предисловие

$$\text{пиздец.}$$

ℵ определение
самое важное в определении
доказательство

```
комментарии от меня
```

$$\sum \int \bigcup \text{ крутые формулы } \iint \prod \bigcap$$

$\sum \int \bigcup$ формулы поменьше $\iint \prod \bigcap$

$\sum \int \bigcup$ ваще маленькие формулы жесть $\iint \prod \bigcap$

итак.

во-первых, земля стекловатой.
во-вторых, далее будет рукописное пособие по тому, как красиво писать греческие буквы, чтобы вы не обосрались на экзамене из-за того, что пишете 3 вместо ξ.
в-третьих, да пребудет с вами Сила.

искренне ваша любительница математических операторов,
$\Sigma o\varphi \acute{\iota} \alpha$

```
а кроме шуток, экзамен длится 1,5-2 часа и состоит из
четырех вопросов; каждый вопрос взят из своей части. в
начале простые, в конце сложные. билетики расписаны из
конспектов, чужих страданий, учебника Бочарова и
Печкина и других сомнительных источников Зименковой С.
22205 в 2023 году.
```

Как писать тупые буквы по-гречески если мы вам лингвисты или кто

краткое пособие

Ω ω омега

осторожнее, δ это дельта

Σ σ сигма σ σ σ но если у вас прямые руки σ σ σ σ (у меня кривые)

σ сигма-алгебра σ (я еще пишу 1. a 2. a так красивее)

ξ или ξ кси (не ξ и не ζ ← это дзета Z ζ)

1. палка
2. ε буква e (ε)
3. ξ хвостик

Рогов пишет её как ξ или ξ но это ξ

H η эта η (не ню; N ν - ню)

M μ μ мю

ʃ + ч = μ так красивее

или

μ так быстрее

1 основные понятия теории вероятностей

из конспекта

$\aleph$ событие — явление, появляющееся при определенных условиях опыта/эксперимента, которые можно воспроизводить неограниченное число раз.
$\aleph$ эксперимент заключается в наблюдении за объектами или явлениями в строго определенных условиях и измерении значений, заранее определенных каким-то признаком этих объектов.

события бывают:

1. достоверные
2. невозможные
3. случайные
    1. совместные и несовместные
    2. зависимые и независимые
    3. противоположные
    4. равновозможные

$\aleph$
событие называется достоверным, если оно обязательно произойдет в результате эксперимента. обозначается Ω.
событие называется невозможным, если оно никогда не произойдет в результате эксперимента. обозначается $\varnothing$.
событие называется случайным, если оно может произойти или не произойти в результате эксперимента. лично я все события помечаю квадратиком: $\square A$.

$\aleph$
два события называются независимыми, если $P(AB) = P(A) \cdot P(B)$.
два события называются зависимыми, если это равенство не выполняется, т. е. P любого события зависит от P других событий.

$\aleph$
два события называются совместными, если могут происходить одновременно.
два события называются несовместными, если появление одного события исключает появление другого события.
события называются равновозможными, если нет оснований считать, что одно из них происходит с большей частотой.

ℵ два события A и $\overline{A}$ называются противоположными, если не появление одного влечет за собой появление другого.
ℵ если группа событий такова, что в результате испытания обязательно должно произойти хотя бы одно из событий, но все остальные в таком случае не происходят, называется полной группой.

ℵ элементарные исходы — исходы испытаний, которые в рамках данной задачи не могут быть разделены на другие исходы.

ℵ элементарные исходы, образующие событие A — благоприятные исходы.

ℵ события A и B называются равными, если наступление $\square A \iff$ наступление $\square B$

ℵ объединение или сумма событий A и B — событие C, которое означает появление хотя бы одного из двух событий A и B.

ℵ разность событий A и B — событие C, которое означает появление события A, но не появление события B.

дополнительная инфа из учебника

ℵ симметрическая разность событий A и B — событие C, которое представляет собой объединение событий $A \setminus B$ и $B \setminus A$:

$$C = A \triangle B = (A \setminus B) \cup (B \setminus A)$$

поскольку события $A \setminus B$ и $B \setminus A$ несовместны, симметрическую разность можно записать также в виде:

$$A \triangle B = (A \setminus B) + (B \setminus A)$$

```
прошу заметить, что U и + это не одно и то же.
```

В том случае, когда события A и B несовместны, наряду со знаком «$\cup$» для их объединения употребляют знак «$+$». Обычно знак «$+$» применяют тогда, когда заведомо известно, что A и B несовместны, и это особо хотят подчеркнуть. В частности, поскольку невозможное событие несовместно с любым событием A, то

$$\varnothing \cup A = \varnothing + A = A.$$

Аналогично определяется объединение трех и более событий. При этом знак «$+$» используется в случае попарной несовместности входящих в объединение событий.

в итоге получается, что симметрическая разность — это объединение событий без их общей части.

ℵ дополнение события A — событие $\overline{A}$, которое происходит тогда и только тогда, когда не происходит событие A. события A и $\overline{A}$ — противоположные.

справедливы формулы:

$$\overline{\Omega} = \varnothing, \quad \overline{\varnothing} = \Omega, \quad \overline{\overline{A}}.$$

$$A \setminus B = A\overline{B}$$

еще у нас есть понятие включения: событие A принадлежит (содержится в, включается в) событию B (записывается $A \subset B$), если появление события A обязательно влечет за собой появление события B (но не наоборот). $AB = A$. кроме того, если $A \subset B$, то

$$A \cup B = B, \quad A \setminus B = \varnothing, \quad A \triangle B = B \setminus A$$

2 классическая вероятность

$\aleph$ вероятность события A — математическая оценка возможности появления этого события в результате эксперимента.

$\aleph$ вероятность события равна $P(A) = \dfrac{M}{N}$, где M — число благоприятных событию A исходов, N — число всех исходов, образующих полную группу событий.

свойства

1. $P(\Omega) = 1$.
   доказательство: $\Omega = A \Rightarrow P(A) = \dfrac{M}{N};\ M = N \Rightarrow P(A) = 1$
2. $P(\varnothing) = 0$.
   доказательство: $\varnothing = A \Rightarrow P(A) = \dfrac{M}{N};\ M = 0 \Rightarrow P(A) = 0$
3. вероятность любого испытания есть неотрицательное число, не превосходящее 1, то есть $0 \leq P(A) \leq 1$.

вероятность — мера случайности.

общая схема решения задач

1. определить, в чем состоит случайный эксперимент и какие у него элементарные события (исходы). убедиться, что элементарные исходы равновозможны.
2. найти общее число элементарных исходов (N).
3. найти число благоприятных для A исходов (M).
4. разделить число благоприятных исходов на число элементарных исходов и получить вероятность события A $\left(P(A) = \dfrac{M}{N}\right)$.

геометрическая вероятность

в некоторой области Ω выбирается точка. какова вероятность, что точка попадет в область A, на прямую L?

$$P(A) = \frac{S(A)}{S(\Omega)},\ \text{где } S(A) \text{ — площадь области } A.$$

если рассматривать прямую L на плоскости, то получится фигня:

$$P(L) = \frac{S(L)}{S(\Omega)} = \frac{0}{S(\Omega)} = 0$$

а все потому, что у прямой площадь равна нулю. то есть вы типа никогда на самом деле не попадете точкой в прямую, только рядышком.
но вообще такая формула работает на числовой прямой и в пространстве:

$$P(A) = \frac{V(A)}{V(\Omega)}$$

$$P(A) = \frac{l(A)}{l(\Omega)}$$

геометрическое определение вероятности

ℵ если предположить, что попадание в любую точку области Ω равновозможно, то вероятность попадания случайной точки в заданное множество A равна

$$P(A) = \frac{S(A)}{S(\Omega)}$$

3 условная вероятность

пусть A и B — события в опыте, $P(B) > 0$

ℵ $P(^A|_B)$ — вероятность события A при условии, что событие B произошло.

$$P(^A|_B) = \frac{P(AB)}{P(B)}$$

следствие:
$P(AB) = P(^A|_B) \cdot P(B) = P(^B|_A) \cdot P(A)$

можно доказать, что таким образом введенная вероятностная мера удовлетворяет всем свойствам вероятностного пространства.

$B = \Omega$
$A = \ ^A|_\Omega$
$$P(A) = P(^A|_\Omega) = \frac{P(A\Omega)}{P(\Omega)} = P(A)$$

свойства

1. $0 \leq P(^A|_B) \leq 1$

2. $A \subset C \quad P(^A|_B) \leq P(^C|_B)$

3. $P(^\Omega|_B) = 1; \quad P(^\varnothing|_B) = 0$

4. пусть A и C несовместные $(A \cdot C = \varnothing)$. тогда:
   $P(^{(A+C)}|_B) = P(^A|_B) + P(^C|_B)$
   док-во: просто подставить в формулу.
   $P((A+C)B) = P(AB + CB) = P(AB) + P(CB)$ и делим на $P(B)$.

5. для любых A и C:
   $P(^{(A+C)}|_B) = P(^A|_B) + P(^C|_B) - P(^{AC}|_B)$
   (если они несовместные, то $P(AC) = 0$, и последнее слагаемое уберется.)

6. $P(^{\bar{A}}|_B) = 1 - P(^A|_B)$
док-во: $(A + \bar{A})B = AB + \bar{A}B = 1$ и делим на $P(B)$.
(или подставить $\bar{A} = 1 - A$ в формулу и посмотреть, что будет.)

7. аксиома полноты: $C \subset A, P(^A|_B) = 0 \implies \exists\, P(^C|_B)$ (???)

8. аксиома счетной аддитивности:
$\{A_i\}_{i \in I},\ A_i \cdot A_j = \varnothing\ при\ i \neq j$ (попарно несовместны)
$P\left(^{\sum\limits_{i \in I} A_i}\Big|_B\right) = \sum\limits_{i \in I} P(^{A_i}|_B)$
доказывается методом мат индукции аналогично с 5: подставить в формулу первое и $n+1$-ое, и дальше все логично.

9. $P(AB) = P(B)P(^A|_B)$ — это следует из определения.
пусть у нас есть $\{A_i\}_{i \in I}^n$, и $P(A_1 \cdot \ \dots\ \cdot A_{n-1}) > 0$ — второе условие нужно, чтобы не делить на ноль.
тогда получается, что мы вот в ту формулу из начала подставляем произведение всех событий сразу:

$$P(\prod_{i=1}^{n} A_i) = P(A_1) \cdot P(^{A_2}|_{A_1}) \cdot P(^{A_3}|_{A_1 A_2}) \cdot\ \dots\ \cdot P(^{A_n}|_{A_1 \cdot \dots \cdot A_{N-1}})$$

доказательство опять методом мат индукции.

2. Формула умножения вероятностей

На практике часто происходит так, что известны или достаточно просто определяются именно условные вероятности и с их помощью необходимо вычислить безусловную вероятность некоторого события. Простейшей формулой для решения задач такого типа является формула умножения вероятностей. Пусть имеются события $A_1, A_2, \ldots, A_n$. Тогда, используя понятие условной вероятности, можно написать

$$\mathsf{P}(A_n \mid A_1A_2\ldots A_{n-1}) = \frac{\mathsf{P}(A_1A_2\ldots A_n)}{\mathsf{P}(A_1A_2\ldots A_{n-1})}$$

или, умножая на $\mathsf{P}(A_1A_2\ldots A_{n-1})$,

$$\mathsf{P}(A_1A_2\ldots A_n) = \mathsf{P}(A_1A_2\ldots A_{n-1}) \cdot \mathsf{P}(A_n \mid A_1A_2\ldots A_{n-1}).$$

Аналогичным образом

$$\mathsf{P}(A_1A_2\ldots A_{n-1}) = \mathsf{P}(A_1A_2\ldots A_{n-2}) \cdot \mathsf{P}(A_{n-1} \mid A_1A_2\ldots A_{n-2}).$$

Продолжая эту процедуру, получаем

$$\mathsf{P}(A_1A_2\ldots A_n) = \mathsf{P}(A_1)\mathsf{P}(A_2|A_1)\mathsf{P}(A_3|A_1A_2)\ldots\mathsf{P}(A_n|A_1A_2\ldots A_{n-1}).$$

Последнее соотношение носит название *формулы умножения вероятностей.*

```
на лекции не доказывали, а на экзамене видимо
придется, но я уже не претендую. :D
```

10. формула полной вероятности:
пусть $H_1,\ H_2,\ \ldots,\ H_K$ — полная группа несовместных событий (гипотез). $H_i \cdot H_j \underset{i\neq j}{=} \varnothing \quad \sum\limits_{i=1}^{k} P(H_i) = 1$
тогда справедлива следующая формула:

$$P(A) = P({}^{A}|_{H_1})P(H_1) + P({}^{A}|_{H_2})P(H_2) + \ \ldots\ + P({}^{A}|_{H_n})P(H_n)$$

это формула полной вероятности.

4 независимые события. принцип независимости

ℵ события A и B называются независимыми (стохастически), если $P(AB) = P(A) \cdot P(B)$.

на самом деле эта формулка — следствие более конкретного определения.
ℵ события A и B называются независимыми, если условная вероятность события B при условии A равна безусловной вероятности события B, т. е. $P({}^B|_A) = P(B)$. это понятие симметрично относительно перестановки событий A и B в формуле, т. е. если событие B не зависит от события A, то и событие A не зависит от события B. для доказательства этого можно подставить $P({}^B|_A) = \dfrac{P(AB)}{P(A)}$ в формулу:

$$P({}^B|_A) = P(B)$$

$$\frac{P(AB)}{P(A)} = P(B)$$

$$P(AB) = P(B) \cdot P(A)$$

$$\frac{P(AB)}{P(B)} = P(A)$$

$$P({}^A|_B) = P(A)$$

так что ту прикольную милую простую понятную формулу из определения мы получаем в ходе доказательства того, что определение с условной вероятностью работает в обе стороны.

свойства

1. пусть A и B — независимые. тогда $(\overline{A},\ B)$, $(A,\ \overline{B})$, $(\overline{A},\ \overline{B})$ — тоже независимые.
   доказательство:

$$P({}^{\overline{A}}|_B) = 1 - P({}^A|_B) = 1 - P(A) = P(\overline{A})$$

со всеми остальными аналогично.

2. события в семействе событий $\{A_i\}_{i\in I}^n$ называются независимыми в совокупности, если для любого $2 \le k \le n$ и любого набора $A_1, \ \dots, \ A_k$ выполняется равенство:

$$P(A_1 \cdot \ \dots \ \cdot A_k) = P(A_1) \cdot \ \dots \ \cdot P(A_k)$$

3. события в семействе событий $\{A_i\}_{i\in I}^n$ называются попарно независимыми, если для любого $i \ne j$ выполняется равенство:

$$P(A_iA_j) = P(A_i) \cdot P(A_j)$$

из независимости в совокупности следует попарная независимость, но из попарной независимости не следует независимость в совокупности.

```
(далее рассматривается пример с тетраэдром. с точки
зрения физики он немножко бредовый, но с точки
зрения теории вероятностей все ок - если
произведение вероятностей равно вероятности
произведения, то события независимы. на то
стохастическая независимость и стохастическая.)
```

Определим теперь понятие независимости событий в совокупности. Назовем события A, B и C *независимыми* (*в совокупности*), если $\mathsf{P}(AB) = \mathsf{P}(A)\,\mathsf{P}(B)$, $\mathsf{P}(AC) = \mathsf{P}(A)\,\mathsf{P}(C)$, $\mathsf{P}(BC) = \mathsf{P}(B)\,\mathsf{P}(C)$ и, кроме того, $\mathsf{P}(ABC) = \mathsf{P}(A)\,\mathsf{P}(B)\,\mathsf{P}(C)$. Аналогично определяется понятие независимости в совокупности и большего числа событий: вероятность пересечения любых двух различных событий равна произведению вероятностей этих событий; вероятность пересечения любых трех событий равна произведению их вероятностей; ...; вероятность пересечения всех событий равна произведению вероятностей. Из независимости событий в совокупности вытекает их попарная независимость. Однако из попарной независимости, вообще говоря, независимость в совокупности не следует, как показывает следующий простой пример (С. Н. Бернштейн).

Пример 11. Три грани правильного тетраэдра раскрашены в синий, красный и зеленый цвета, а в раскраске четвертой грани присутствуют все эти цвета. Событие A — тетраэдр упал на грань, в раскраске которой присутствует синий, событие B — красный и событие C — зеленый цвет. Поскольку каждый цвет присутствует в раскраске двух граней из четырех, то $\mathsf{P}(A) = \mathsf{P}(B) = \mathsf{P}(C) = 1/2$. Далее, два цвета присутствуют в раскраске одной грани. Поэтому $\mathsf{P}(AB) = \mathsf{P}(BC) = \mathsf{P}(AC) = 1/4$, и события A, B и C попарно независимы. Однако все три цвета присутствуют в раскраске также только одной грани и, значит, $\mathsf{P}(ABC) = 1/4 \neq \mathsf{P}(A)\,\mathsf{P}(B)\,\mathsf{P}(C) = 1/8$. Таким образом, хотя события A, B и C попарно независимы, они зависимы в совокупности. □

доп. материал

формула для вероятности объединения независимых событий:
пусть $A = A_1 \cup \ \ldots \ \cup A_n$.
тогда в соответствии с формулой де Моргана $\overline{A} = \overline{A}_1 \cdot \ \ldots \ \cdot \overline{A}_n$.
если события $A_1, \ \ldots, \ A_n$ независимы, то и события $\overline{A}_1, \ \ldots, \ \overline{A}_n$ тоже независимы. тогда получается, что $P(\overline{A}) = P(\overline{A}_1) \cdot \ \ldots \ \cdot P(\overline{A}_n)$.
подставляем сюда $\overline{A} = 1 - A$ и получаем:

$$P(A) = 1 - (1 - P(A_1)) \cdot \ \ldots \ \cdot (1 - P(A_n))$$

принцип независимости

```
часть источников утверждает, что физическая
независимость => стохастическая независимость, но не
наоборот. впрочем, на лекциях утверждалось, что так и
есть, а значит попытаемся смириться.
```

Предположение о стохастической независимости рандомизированных выборов игроков базируется прежде всего на физической независимости выборов. В книге М. Каца [4] отмечается, что, вообще говоря, стохастическая независимость испытаний не вытекает из их физической независимости. Аналогичное рассуждение см. в статье Н. Н. Воробьева [5].

(источник: В. Л. Крепс, О теоретико-игровой характеризации стохастической независимости, Дискрет. матем., 2010, том 22, выпуск 1, 115-125)

принцип независимости: **если события физически независимы, то они также стохастически независимы; обратное утверждение неверно**.

цитируя многоуважаемую Марию, на лекции было следующее: стохастическая независимость есть удобное модельное предположение, охватывающее широкий круг ситуаций. физическая же независимость это что-то очень желаемое, но не всегда достижимое: часто, даже если кажется, что события независимы физически, они на самом деле могут влиять друг на друга. например, если в одном городе позавчера шел дождь, а в другом он пошел сегодня, то нельзя с точностью сказать, что эти два события независимы — циклон мог пригнать тучи из одного места в другое. но чтобы не мучиться, можно объявить стохастическую независимость и небольшой зависимостью в формулах пренебречь.

получается, что:

- события физически независимы $\Rightarrow$ события независимы стохастически;
- события физически зависимы $\not\Rightarrow$ события зависимы стохастически;
- события стохастически независимы $\not\Rightarrow$ события независимы физически;

- события стохастически зависимы $\Rightarrow$ события зависимы физически.

пример

пусть есть колода из 52 карт, из которой случайно достают одну.
☐A — карта будет масти бубны.
☐B — карта будет тузом.

давайте проверим стохастическую независимость:
$P(AB) = \frac{1}{52}$ (ведь в колоде только один туз бубен.)
$$P(A) \cdot P(B) = \underbrace{\frac{1}{4}}_{\text{масть}} \cdot \underbrace{\frac{4}{52}}_{\text{тузы}} = \frac{1}{52}$$
получается, что события A и B независимы стохастически, но вполне могут быть зависимы физически.
теперь добавим в колоду джокера, карт станет 53.
$P(AB) = \frac{1}{53}$
$$P(A) \cdot P(B) = \underbrace{\frac{13}{53}}_{\text{масть}} \cdot \underbrace{\frac{4}{53}}_{\text{тузы}} = \frac{52}{2809} \neq \frac{1}{53}$$
получается, что события не являются стохастически независимыми, а значит и физически они зависимы.

(для понимания физической зависимости в этой ситуации: если мы вытянем джокера, то он уже точно не сможет быть ни тузом, ни бубнами, а значит, эти события зависимы физически.)

5 формула полной вероятности

пусть $H_1, H_2, \ldots, H_n$ — полная группа несовместных событий:

1. $H_i H_j = \varnothing$ при $j \neq j$
2. $H_1 + H_2 + \ \ldots + H_n = \Omega$

события, удовлетворяющие этим условиям, будем называть гипотезами.

пусть также имеется событие A, и нам известны $P(H_1), \ \ldots, \ P(H_n)$, а также $P(^A|_{H_1}), \ \ldots, \ P(^A|_{H_n})$.

нужно найти безусловную вероятность $P(A)$. для этого представим A в следующем виде:

$$A = A\Omega = A(H_1 + \ \ldots \ + H_n) = AH_1 + \ \ldots \ + AH_n$$

(можем себе позволить такое извращение, потому что гипотезы несовместны.)
тогда

$$P(A) = P(AH_1) + \ \ldots \ + P(AH_n)$$

по формуле умножения вероятностей:

$$P(AH_1) = P(H_1)P(^A|_{H_1}), \ \ldots, \ P(AH_n) = P(H_n)P(^A|_{H_n})$$

подставляем это в предыдущую формулу и получаем такую, в которой нам все известно:

$$P(A) = P(H_1)P(^A|_{H_1}) + \ \ldots \ + P(H_n)P(^A|_{H_n})$$

это формула полной вероятности.

Пример 13. Частица пролетает мимо трех счетчиков, причем она может попасть в каждый из них с вероятностями 0,3, 0,2 и 0,4. В свою очередь, если частица попадает в первый счетчик, то она регистрируется с вероятностью 0,6, во второй — с вероятностью 0,5 и в третий — с вероятностью 0,55. Найдем вероятность того, что частица будет зарегистрирована (событие A). В нашем случае гипотеза H_1 — частица попадает в первый счетчик ($\mathsf{P}(H_1) = 0{,}3$), H_2 — во второй ($\mathsf{P}(H_2) = 0{,}2$) и H_3 — в третий ($\mathsf{P}(H_3) = 0{,}4$). События H_1, H_2 и H_3 не пересекаются, однако они не составляют полной группы событий. Для того чтобы получить полную группу событий, нужно добавить событие H_4, заключающееся в том, что частица не попадет ни в один счетчик. Ясно, что $\mathsf{P}(H_4) = 1 - \mathsf{P}(H_1) - \mathsf{P}(H_2) - \mathsf{P}(H_3) = 1 - 0{,}3 - 0{,}2 - 0{,}4 = 0{,}1$. Условные вероятности события A при условии каждой гипотезы равны: $\mathsf{P}(A \mid H_1) = 0{,}6$, $\mathsf{P}(A \mid H_2) = 0{,}5$, $\mathsf{P}(A \mid H_3) = 0{,}55$ и $\mathsf{P}(A \mid H_4) = 0$. По формуле полной вероятности имеем

$$\mathsf{P}(A) = 0{,}3 \cdot 0{,}6 + 0{,}2 \cdot 0{,}5 + 0{,}4 \cdot 0{,}55 + 0{,}1 \cdot 0 = 0{,}5. \qquad \square$$

6 формула Байеса

пусть нам известны полная группа событий $H_1, \ldots, H_n$ $\left(H_iH_j = \varnothing, \sum\limits_{i=1}^{n} H_i = \Omega\right)$, известны условные вероятности события A при условиях $H_1, \ldots, H_n$. нужно найти $P(^{H_i}|_A)$.

$P(^{H_i}|_A) = \dfrac{P(AH_i)}{P(A)}$; $\quad P(AH_i) = P(H_i)P(^A|_{H_i})$.

тогда:

$$P\left(^{H_i}|_A\right) = \frac{P(H_i)P\left(^A|_{H_i}\right)}{P(A)}$$

это выражение называется формулой Байеса.

подставляя вместо $P(A)$ формулу полной вероятности, мы получаем чуть более *полную* (поняли каламбур?) формулу:

$$P\left(^{H_i}|_A\right) = \frac{P(H_i)P\left(^A|_{H_i}\right)}{\sum\limits_{i=1}^{n} P(H_i)P(^A|_{H_i})}$$

Пример 14. Три завода выпускают одинаковые изделия, причем первый завод производит 50%, второй — 20% и третий — 30% всей продукции. Первый завод выпускает 1% брака, второй завод — 8% и третий — 3%. Наудачу выбранное изделие оказалось бракованным (событие A). Найдем вероятность того, что оно изготовлено на втором заводе. У нас имеется три гипотезы: H_1 — изделие изготовлено на первом заводе, H_2 — на втором заводе и H_3 — на третьем. По условию задачи $\mathsf{P}(H_1) = 0{,}5$, $\mathsf{P}(H_2) = 0{,}2$, $\mathsf{P}(H_3) = 0{,}3$, $\mathsf{P}(A \mid H_1) = 0{,}01$, $\mathsf{P}(A \mid H_2) = 0{,}08$, $\mathsf{P}(A \mid H_3) = 0{,}03$. Условная вероятность того, что бракованное изделие изготовлено на втором заводе, определяется формулой Байеса:

$$\mathsf{P}(H_2 \mid A) = \frac{\mathsf{P}(H_2)\mathsf{P}(A \mid H_2)}{\mathsf{P}(H_1)\mathsf{P}(A \mid H_1) + \mathsf{P}(H_2)\mathsf{P}(A \mid H_2) + \mathsf{P}(H_3)\mathsf{P}(A \mid H_3)} =$$

$$= \frac{0{,}2 \cdot 0{,}08}{0{,}5 \cdot 0{,}01 + 0{,}2 \cdot 0{,}08 + 0{,}3 \cdot 0{,}03} = \frac{8}{15}.$$

Итак, несмотря на то что продукция второго завода составляет $1/5$, его доля в браке больше половины. $\square$

7 формула умножения вероятностей для n событий

пусть $A_1, \ldots, A_n$ — случайные события.
$\{A_i\}_{i=1}^n, \quad P(A_1 \cdot \ldots \cdot A_{n-1}) > 0$

$$P\left(\prod_{i=1}^{n} A_i\right) = P(A_1) \cdot P\left({}^{A_2}|_{A_1}\right) \cdot P\left({}^{A_3}|_{A_1 A_2}\right) \cdot \ \ldots \ \cdot P\left({}^{A_n}|_{A_1 \ldots A_{n-1}}\right)$$

доказательство

доказывается методом мат индукции.
это известно по формуле условной вероятности: $P(AB) = P(B) \cdot P({}^{A}|_{B})$, и это же есть доказываемая нами формула для наименьшего n. значит, теперь нам надо доказать, что если это верно для $n-1$, то это также верно для n, и все.

пусть утверждение верно для $n-1$:

$$P\left(\prod_{i=1}^{n-1} A_i\right) = P(A_1) \cdot P({}^{A_2}|_{A_1}) \cdot \ \ldots \ \cdot P({}^{A_{n-1}}|_{A_1 \ldots A_{n-2}})$$

пусть $\prod_{i=1}^{n-1} A_i = B$. тогда $P\left(\prod_{i=1}^{n-1} A_i\right) = P(B)$.

$\left(\prod_{i=1}^{n-1} A_i\right) \cdot A_n = \left(\prod_{i=1}^{n} A_i\right) \Rightarrow A_n B = \left(\prod_{i=1}^{n} A_i\right)$

подставим это в формулу в начале:

$P(A_n B) = P\left(\prod_{i=1}^{n-1} A_i\right) \cdot P({}^{A_n}|_{B}) =$
$P(A_1) \cdot P({}^{A_2}|_{A_1}) \cdot \ \ldots \ \cdot P({}^{A_{n-1}}|_{A_1 \ldots A_{n-2}}) \cdot P({}^{A_n}|_{A_1, \ldots, A_{n-1}}) \Rightarrow$
$P\left(\prod_{i=1}^{n} A_i\right) = P(A_1) \cdot P\left({}^{A_2}|_{A_1}\right) \cdot P\left({}^{A_3}|_{A_1 A_2}\right) \cdot \ \ldots \ \cdot P\left({}^{A_n}|_{A_1 \ldots A_{n-1}}\right)$, что и требовалось доказать.

в случае, если события $A_1, \ \ldots, \ A_n$ — независимые в совокупности, то $P\left(\prod\limits_{i=1}^{n} A_i\right) = \prod\limits_{i=1}^{n} P\left(A_i\right)$.

8 основные соотношения между вероятностями событий

пусть задано вероятностное пространство $(\Omega, \mathfrak{A}, P)$. (здесь если что $\mathfrak{A}$ это А готическое, σ-алгебра.)
событие A — подмножество пространства элементарных исходов Ω, принадлежащее σ-алгебре $\mathfrak{A}$. каждому событию A поставлено в соответствие число $P(A)$.

числовая функция $P(A)$, заданная на σ-алгебре $\mathfrak{A}$, называется вероятностью события A, если она удовлетворяет следующим аксиомам:

1. $P(A) \geq 0$ (аксиома неотрицательности)
2. $P(\Omega) = 1$ (аксиома нормированности)
3. $P(A+B) = P(A) + P(B)$ (аксиома сложения, также называют аксиомой счетной аддитивности), если $A,\ B \in \mathfrak{A}$ и $A \cup B = \varnothing$ (т. е. они несовместные).
4. $P(A) = 0 \Rightarrow \forall\ B \subset A\ P(B) = 0$ (аксиома полноты)
   если вероятность события A равна нулю, то вероятности его составляющих тоже равна нулю.

более полная формулировка аксиомы сложения

$\{A_i\}_{i \in I}$, где I — конечное/счетное множество. $\forall i,\ j \in I \quad i \neq j \quad A_i \cdot A_j = \varnothing$.
тогда $P(\bigcup\limits_{i \in I} A_i) = \sum\limits_{i \in I} P(A_i)$.

основные соотношения между вероятностями событий.

1. вероятность дополнительного события

$A \in \mathfrak{A},\ \overline{A} = \Omega \setminus A$.
тогда

$$P(\overline{A}) = 1 - P(A)$$

доказательство:

$A \cdot \overline{A} = \varnothing$ (т. к. эти события несовместны.)
$P(\Omega) = P(A + \overline{A}) = P(A) + P(\overline{A})$ (по аксиоме сложения и аксиоме нормированности.)

2. большему событию соответствует большая вероятность

даны события A и B. $A \subset B$. тогда

$$P(A) \leq P(B)$$

доказательство:
$B = (B \setminus A) + A$
по аксиоме сложения:
$P(B) = \underbrace{P(B \setminus A)}_{\geq 0, \text{ аксиома } 1} + P(A)$
$P(B \setminus A) \geq 0$ по аксиоме неотрицательности, а значит, $P(B) \geq P(A)$.

3. вероятность заключена между 0 и 1

$$0 \leq P(A) \leq 1$$

доказательство: $A \subset \Omega$ по определению события A, по аксиоме нормированности $P(\Omega) = 1$, по соотношению 2 (предыдущему) $P(A) \leq P(\Omega) \Rightarrow P(A) \leq 1$. в совокупности с аксиомой неотрицательности $(P(A) \geq 0)$ получается, что $0 \leq P(A) \leq 1$.

4. вероятность невозможного события

$$P(\varnothing) = 0$$

доказательство:

1. как в учебнике:
   $A = A + \varnothing \Rightarrow$ (по аксиоме сложения) $P(A) = P(A) + P(\varnothing) \Rightarrow$
   $P(\varnothing) = P(A) - P(A) = 0$.
2. как на лекции было:
   (по соотношению 1) $P(\overline{\Omega}) = 1 - P(\Omega) = 1 - 1 = 0$.

5. вероятности объединения и разности

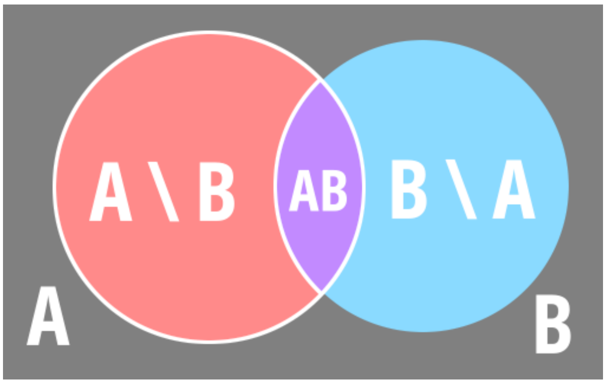
$$P(A \setminus B) = P(A) - P(AB) \qquad (1)$$

$$P(B \setminus A) = P(B) - P(AB) \tag{2}$$

$$P(A + B) = P(A) + P(B) - P(AB) \tag{3}$$

доказательство:

1. $A = (A \setminus B) + AB \Rightarrow P(A) = P(A \setminus B) + P(AB) \Rightarrow$
   $P(A \setminus B) = P(A) - P(AB)$



2. $B = (B \setminus A) + AB \Rightarrow P(B) = P(B \setminus A) + P(AB) \Rightarrow$
   $P(B \setminus A) = P(B) - P(AB)$

3. $A + B = (A \setminus B) + (B \setminus A) + AB$
   $P(A + B) = P(A \setminus B) + P(B \setminus A) + P(AB)$

   подставляем сюда выражения $P(A \setminus B) = P(A) - P(AB)$ и $P(B \setminus A) = P(B) - P(AB)$.

   получаем:
   $P(A + B) = P(A) - P(AB) + P(B) - P(AB) + P(AB)$
   $P(A + B) = P(A) + P(B) - P(AB)$

6. вероятность объединения событий не больше, чем сумма вероятностей этих событий

пусть $B = \bigcup\limits_{i \in I} A_i$, где I — конечное или счетное множество. тогда

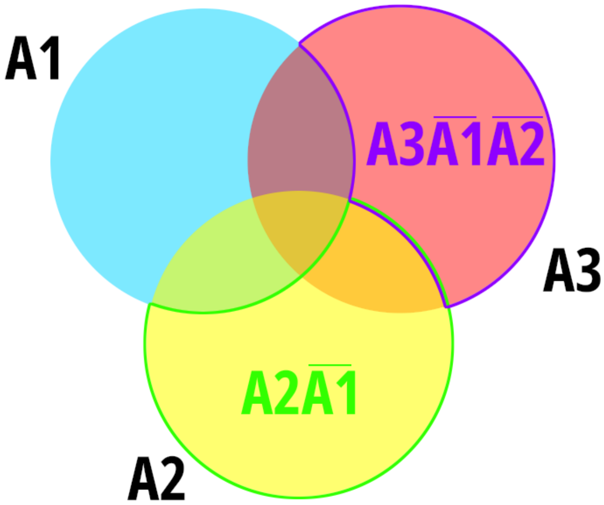
$$P(B) \leq \sum_{i \in I} P(A_i)$$

доказательство:
$B = A_1 + A_2\overline{A}_1 + A_3\overline{A}_1\overline{A}_2 + \ \ldots$

$P(B) = P(A_1) + \underbrace{P(A_2\overline{A}_1)}_{\leq P(A_2)} + \underbrace{P(A_3\overline{A}_1\overline{A}_2)}_{\leq P(A_3)} + \ \dots$

$P(B) \leq P(A_1) + P(A_2) + P(A_3) + \ \dots$

```
по картинке должно быть хорошо видно, почему B это
сумма таких прикольных слагаемых, и почему прикольные
слагаемые меньше или равны чем просто A_i.
```



7. формула суммы n событий

$B = \bigcup\limits_{i=1}^{n} A_i$

$P_1 = \sum\limits_{i=1}^{n} P(A_i)$

$P_2 = \sum\limits_{1 \leq i < j \leq n} P(A_iA_j)$ — все пары по одному разу.

$P_3 = \sum\limits_{1 \leq i < j < k \leq n} P(A_iA_jA_k)$ — все тройки по одному разу.

$...$

$P_n = P(A_1A_2 \dots \ A_n)$

$$P(B) = P_1 - P_2 + P_3 - P_4 + \ \dots \ \pm P_n$$

Последнее свойство допускает очевидное, но весьма полезное обобщение на случай произвольного числа слагаемых

6. $\mathsf{P}(A_1 \cup \ldots \cup A_n) = \mathsf{P}(A_1) + \ldots + \mathsf{P}(A_n) - \mathsf{P}(A_1 A_2) - \mathsf{P}(A_1 A_3) - \ldots$
$\ldots - \mathsf{P}(A_{n-1} A_n) + \mathsf{P}(A_1 A_2 A_3) + \ldots + (-1)^{n+1} \mathsf{P}(A_1 A_2 \ldots A_n).$

Свойство 6 доказывается индукцией по n. Так, для трех событий A, B и C

$$\mathsf{P}(A \cup B \cup C) = \mathsf{P}(A) + \mathsf{P}(B \cup C) - \mathsf{P}(A(B \cup C)) =$$
$$= \mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C) - \mathsf{P}(BC) - \mathsf{P}(AB \cup BC) =$$
$$= \mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C) - \mathsf{P}(BC) - \mathsf{P}(AB) - \mathsf{P}(AC) + \mathsf{P}(ABC).$$

доказательство методом математической индукции:

для $n = 2$: по 5 соотношению.
докажем, что если верно для $n-1$, то верно для n.
$P(\bigcup_{i=2}^n A_i) = \sum_{i=2}^n P(A_i) - \sum_{2 \le i < j \le n} P(A_i A_j) + \ \ldots$
$P(\bigcup_{i=2}^n A_i A_j) = \sum_{i=2}^n P(A_i A_j) - \sum_{2 \le i < j \le n} P(A_i A_j A_k) + \ \ldots$
(это нужно чтобы просто в формулу подставить)

$P(B) = P(A_1 + \bigcup_{i=2}^n A_i) = P(A_1) + P(\bigcup_{i=2}^n A_i) - P(\bigcup_{i=2}^n A_1 A_i)$
$= \sum_{i=1}^n P(A_i) - \sum_{i=1 \le j \le n} P(A_i A_j) + \ \ldots$
красота получилась? красота получилась. но, честно признаться, я ничего не поняла.

9 схема Бернулли (биномиальная схема)

что вообще происходит

опыт в схеме Бернулли состоит из n-кратного повторения какого-то опыта, в каждом из которых с вероятностью p может произойти какое-то событие (будем называть это успехом). тогда с вероятностью $q = 1 - p$ это событие может не наступить (будем называть это неудачей). тогда результат каждого опыта можно записать в виде последовательности 0 и 1, где 0 — неудача, а 1 — успех. типичным представителем схемы Бернулли является подбрасывание монетки, где успех это выпадение одной стороны, а неудача — выпадение другой стороны.

теперь более формально

определим вероятностные пространства $M_k = (\Omega_k,\ \mathfrak{A}_k,\ P_k)$, где $k = \overline{1,\ n}$.
$\Omega = \{\Omega_1;\ \Omega_2;\ \ldots;\ \Omega_n\}$
$\omega \in \Omega$, $\omega = (\omega_1,\ \ldots,\ \omega_n)$, где $\omega_i \in \Omega_i$

событие $A \in \mathfrak{A}$, $A = \{A_1;\ A_2;\ A_3;\ \ldots;\ A_n\}$, $A_i \in \mathfrak{A}_i$.
$P(A) = \sum\limits_{\omega \in A} P(\omega) = \sum\limits_{\omega=(\omega_1,\ \ldots,\ \omega_n)} P_1(\omega_1) \cdot\ \ldots\ \cdot P_n(\omega_n)$

схема Бернулли

$M_k = (\Omega_k,\ \mathfrak{A}_k,\ P_k), \quad k = 1,\ \ldots,\ n$
$\Omega_k = \{0;\ 1\}$, где 1 — успех. (здесь Ω называется бинарным вектором, т. к. принимает значения 0 или 1.)
$P_k(0) = 1 - p$
$P_k(1) = p$

условия:

1. последовательность из n независимых испытаний;
2. в результате k испытаний мы получаем 1 или 0;
3. вероятность успеха не меняется от опыта к опыту.

$\mu(n)$ — число единиц (успехов) среди элементарных исходов $\omega \in \Omega$.

$$P(\mu(n) = k) = C_n^k p^k (1-p)^{n-k}$$

словами: вероятность того, что из n случаев k будут успехами.

пример

с вероятностью $p = 0.8$ стрелок попадет в мишень. стрелок выстрелил в мишень 6 раз. найти вероятность того, что 4 из 6 раз стрелок попадет в мишень.

$$P_6(4) = C_6^4 \cdot 0.8^4 \cdot 0.2^2$$

замечательная табличка, которую нам давали на практиках:

	без упорядочивания	с упорядочиванием
без возвращения	$C_n^k = \dfrac{n!}{k!(n-k)!}$	$A_n^k = \dfrac{n!}{(n-k)!}$
с возвращением	$C_{n+k-1}^k = \dfrac{(n+k-1)!}{k!(n-1)!}$	n^k

считаем:

$P_6(4) = \frac{6!}{4! \cdot 2!} \cdot 0.4096 \cdot 0.04 = 15 \cdot 0.016384 = 0.24576 \approx 0.25$

10 полиномиальная схема

простыми словами, как в учебнике

полиномиальная схема — обобщение схемы Бернулли. в n испытаниях может произойти одно и только одно из m несовместных событий $A_1, \ldots, A_m$, причем событие A_i наступает с вероятностью p_i. тогда вероятность $P(n_1, \ldots, n_m)$ того, чьл в n испытаниях событие A_1 произойдет ровно n_1 раз, ..., событие A_m произойдет ровно n_m раз, определяется выражением:

$$P(n_1, \ldots n_m) = \frac{n!}{n_1! \cdot \ \ldots \ \cdot n_m!} \cdot p_1^{n_1} \cdot \ \ldots \ \cdot p_m^{n_m}$$

абсолютная гадость на математическом языке

$\Omega_n = \{1, 2, \ldots, m\}$ (в каждом из n опытов может выпасть такой-то по номеру результат.)

$P_n(i) = p_i, \ i = 1, \ldots, m$ (в каждом из n опытов вероятность i-того результата равна p_i.)

$\sum\limits_{i=1}^{m} p_i = 1$ (события несовместны, и если сложить все вероятности исходов, получится 1.)

$\omega:$ (элементарный исход)

"1" — n_1 раз;

"2" — n_2 раз;

$\ldots$

"m" — n_m раз.

$\sum\limits_{i=1}^{m} n_i = n$ (всего провели n опытов, и в каждом выпал какой-то результат, поэтому если количество всех результатов сложить, то получится n.)

тогда:

$$P(\text{"1"} - n_1, \text{"2"} - n_2, \ldots, \text{"m"} - n_m) = \frac{n!}{n_1! \cdot \ \ldots \ \cdot n_m!} \cdot p_1^{n_1} \cdot \ \ldots \ \cdot p_m^{n_m}$$

11 формула сложения и умножения вероятностей

формула сложения вероятностей

для несовместных событий

$$P(A+B) = P(A) + P(B)$$

для совместных событий

$$P(A+B) = P(A) + P(B) - P(AB)$$

для n произвольных событий

$B = \bigcup\limits_{i=1}^{n} A_i$

$P_1 = \sum\limits_{i=1}^{n} P(A_i)$

$P_2 = \sum\limits_{1\le i<j\le n} P(A_iA_j)$ — все пары по одному разу.

$P_3 = \sum\limits_{1\le i<j<k\le n} P(A_iA_jA_k)$ — все тройки по одному разу.

$\ldots$

$P_n = P(A_1A_2\ldots A_n)$

$$P(B) = P_1 - P_2 + P_3 - P_4 + \ \ldots \ \pm P_n$$

Последнее свойство допускает очевидное, но весьма полезное обобщение на случай произвольного числа слагаемых

6. $\mathsf{P}(A_1 \cup \ldots \cup A_n) = \mathsf{P}(A_1) + \ldots + \mathsf{P}(A_n) - \mathsf{P}(A_1A_2) - \mathsf{P}(A_1A_3) - \ldots$
$\ldots - \mathsf{P}(A_{n-1}A_n) + \mathsf{P}(A_1A_2A_3) + \ldots + (-1)^{n+1}\mathsf{P}(A_1A_2\ldots A_n).$

Свойство 6 доказывается индукцией по n. Так, для трех событий A, B и C

$$\begin{aligned}\mathsf{P}(A\cup B\cup C) &= \mathsf{P}(A) + \mathsf{P}(B\cup C) - \mathsf{P}(A(B\cup C)) = \\ &= \mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C) - \mathsf{P}(BC) - \mathsf{P}(AB\cup BC) = \\ &= \mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C) - \mathsf{P}(BC) - \mathsf{P}(AB) - \mathsf{P}(AC) + \mathsf{P}(ABC).\end{aligned}$$

формула умножения вероятностей

для независимых событий

$$P(AB) = P(A)\cdot P(B)$$

для зависимых событий

$$P(AB) = P(A) \cdot P(^{B}|_{A}) = P(B) \cdot P(^{A}|_{B})$$

для n произвольных событий

$$P\left(A_1 A_2 \ldots A_n\right) = P\left(A_1\right) \cdot P\left(^{A_2}|_{A_1}\right) \cdot P\left(^{A_3}|_{A_1 A_2}\right) \cdot \ \ldots \ \cdot P\left(^{A_n}|_{A_1 A_2 \ldots A_{n-1}}\right)$$

12 аксиоматическая теория вероятностей

Ω — множество элементарных исходов. может быть любым множеством.
$\mathfrak{A}$ $\mathfrak{a}$ (\mathfrak A) — "a" готическое.
элементы подмножества $\mathfrak{A}$ — события.

```
это та фигня, которой нас на лекциях мучили, мол случайное
событие — не событие, которое может произойти или не
произойти, а элемент сигма-алгебры.
```

ℵ $\mathfrak{A}$ — алгебра, если $\varnothing \in \mathfrak{A}$, $\Omega \in \mathfrak{A}$ и $A,\ B \in \mathfrak{A} \to AB \in \mathfrak{A},\ A + B \in \mathfrak{A},\ A \setminus B \in \mathfrak{A}$.
(невозможное и достоверное события принадлежат алгебре, и если события A и B принадлежат алгебре, то их объединение и разность тоже принадлежат алгебре.)

ℵ $\mathfrak{A}$ это σ-алгебра (сигма-алгебра), если $\{A_i\}_{i\in I}\ \prod_{i\in I} A_i \in \mathfrak{A},\ \sum_{i\in I} A_i \in \mathfrak{A}$.
(для любого набора событий, пересечения и объединения этих событий тоже принадлежат σ-алгебре.)

если $A \in \mathfrak{A}$, определим такую вероятностную меру $P(A)$, что:

1. $P(A) \geq 0$ (аксиома неотрицательности)
2. $P(\Omega) = 1.$ (аксиома нормированности)
3. $\{A_i\}_{i\in I}$, где I — конечное/счетное множество. $\forall i,\ j \in I \quad i \neq j \quad A_i \cdot A_j = \varnothing$.
   тогда $P(\bigcup\limits_{i\in I} A_i) = \sum\limits_{i\in I} P(A_i)$. (аксиома сложения, еще называются аксиомой счетной аддитивности)
4. если $A \in \mathfrak{A};\ P(A) = 0 \quad \Rightarrow \quad \forall B \subset A \quad P(B) = 0$ (аксиома полноты)

ℵ $(\Omega,\ \mathfrak{A},\ P)$ — вероятностное пространство.
ℵ элемент σ-алгебры $\mathfrak{A}$ — случайное событие.

замечание

можно доказать, что классическое определение вероятности и геометрическая вероятность удовлетворяют аксиомам вероятностного пространства.

13 схема выбора, приводящая к сочетаниям

система Уоррена

	без упорядочивания	с упорядочиванием
без возвращения	$C_n^k = \dfrac{n!}{k!(n-k)!}$	$A_n^k = \dfrac{n!}{(n-k)!}$
с возвращением	$C_{n+k-1}^k = \dfrac{(n+k-1)!}{k!(n-1)!}$	n^k

схема выбора, приводящая к сочетаниям

если опыт состоит в выборе k-элементного подмножества из n-элементного множества без упорядочивания и возвращения, то количество исходов считается как число сочетаний $C_n^k = \dfrac{n!}{k!(n-k)!}$.

14 схема выбора, приводящая к размещениям

система Уоррена

	без упорядочивания	с упорядочиванием
без возвращения	$C_n^k = \dfrac{n!}{k!(n-k)!}$	$A_n^k = \dfrac{n!}{(n-k)!}$
с возвращением	$C_{n+k-1}^k = \dfrac{(n+k-1)!}{k!(n-1)!}$	n^k

схема выбора, приводящая к размещениям

если опыт состоит в выборе k-элементного подмножества из n-элементного подмножества без возвращения, но с упорядочиванием, то количество исходов считается как количество размещений $A_n^k = \frac{n!}{(n-k)!}$.

15 схема выбора, с возвращением и без упорядочивания

система Уоррена

	без упорядочивания	с упорядочиванием
без возвращения	$C_n^k = \dfrac{n!}{k!(n-k)!}$	$A_n^k = \dfrac{n!}{(n-k)!}$
с возвращением	$C_{n+k-1}^k = \dfrac{(n+k-1)!}{k!(n-1)!}$	n^k

схема выбора, приводящего к сочетаниям (без упорядочивания) с возвращением

если опыт состоит в выборе k-элементного подмножества из n-элементного множества без упорядочивания, но с возвращением, то количество исходов считается как число сочетаний $C_{n+k-1}^k = \dfrac{(n+k-1)!}{k!(n-1)!}$.

16 дискретная случайная величина, определение и закон распределения

для начала о том, что вообще такое случайная величина

рассмотрим вероятностное пространство $(\Omega,\ \mathfrak{A},\ P)$.
ℵ случайная величина ξ — функция, ставящая в соответствие каждому элементарному исходу ω число $\xi = \xi(\omega)$.
для того, чтобы такое определение было математически корректным, необходимо добавить следующее требование:
для любого числа x множество $\{\omega : \xi(\omega) < x\}$ элементарных исходов ω, для которых $\xi(\omega) < x$, является событием, или, иными словами, принадлежит σ-алгебре $\mathfrak{A}$. это свойство называется измеримостью функции $\xi = \xi(\omega)$ относительно σ-алгебры $\mathfrak{A}$.

получается, что случайная величина — это просто какая-то функция, заданная на пространстве элементарных исходов Ω и измеримая относительно σ-алгебры $\mathfrak{A}$. случайные величины принято обозначать греческими буквами $(\xi,\ \eta,\ \mu)$ и при необходимости добавлять индексы. также для краткости вместо $\{\omega : \xi(\omega) < x\}$ пишут $\{\xi(\omega) < x\}$, если необходимо подчеркнуть связь случайной величины с пространством элементарных исходов Ω, или даже просто $\{\xi < x\}$, если ничего подчеркивать не надо.

ℵ функция распределения случайной величины ξ — такая функция $F(x)$, значение которой в точке x равно вероятности события $\{\xi < x\}$, т. е. события, состоящего только из тех элементарных исходов ω, для которых $\xi(\omega) < x$:

$$F(x) = P(\xi < x)$$

свойства:

1. $0 \le F(x) \le 1$ (т. к. функция распределения это вероятность).
2. $F(x)$ — неубывающая функция (т. е. $F(x_1) \le F(x_2)$ при $x_1 < x_2$).
3. $F(-\infty) = 0, \quad F(+\infty) = 1$ (т. к. $\{\xi < -\infty\} = \varnothing$ и $\{\xi < +\infty\} = \Omega$)

4. $P(x_1 \le \xi < x_2) = F(x_2) - F(x_1)$ (т. к. событие $\{\xi < x_2\}$ при $x_1 < x_2$ представляет собой объединение двух непересекающихся событий: $\{\xi < x_1\}$ — случайная величина приняла значение, меньшее x_1, и $\{x_1 \le \xi < x_2\}$ — случайная величина приняла значение в интервале $[x_1, x_2)$. свойство получаем из аксиомы сложения.)
5. $F(x) = F(x-0)$, т. е. $F(x)$ — непрерывная слева.
   док-во: пусть $x_1,\ \dots,\ x_n,\ \dots$ — возрастающая последовательность чисел, сходящаяся к x. тогда событие $A = \{\xi < x\}$ является счетным объединением несовместных событий

   $$A_1 = \{\xi < x_1\}, \quad A_n = \{x_{n-1} \le \xi < x_n\} \quad (n = 2,\ 3,\ \dots),$$

   то есть $A = A_1 + \ \dots\ + A_n + \ \dots$
   по аксиоме сложения $P(A) = P(A_1) + P(A_2) + \ \dots\ + P(A_n) + \ \dots$
   $= F(x_1) + (F(x_2) - F(x_1)) + \ \dots\ + (F(x_n) - F(x_{n-1})) + \ \dots = \lim\limits_{n\to\infty} F(x_n)$.
   следовательно, $F(x)$ — непрерывная слева.

```
ну на пятерку придется рассказать. дальше я часть информации повторю, но считайте просто инфу до черты за необязательный материал, а после черты - за обязательный.
```

рассмотрим дискретное вероятностное пространство: множество элементарных исходов Ω — дискретное, т. е. счетное или конечное.

$(\Omega,\ \mathfrak{A},\ P)$
$\Omega = \{\omega_i\}_{i\in I}$
$\{X_i\}_{i\in I}$ — возможные значения случайной величины.
$\xi(\omega_i)$ — случайная величина.

ℵ дискретной называется случайная величина, которая каждому элементарному исходу ω ставит в соответствие одно из конечного (в общем случае счетного) набора чисел $X_1,\ X_2,\ \dots,\ X_n$.

ℵ $\{X_i,\ P_i\}_{i\in I}$ — закон распределения дискретной случайной величины. закон распределения характеризует случайную величину. его также можно представить в виде таблицы.

вот определение из учебника:
ℵ рядом распределения (вероятностей) случайной величины называется таблица, состоящая из двух строк: в верхней строке перечислены все возможные значения случайной величины, а в нижней — вероятности $p_i = P\{\xi = X_i\}$ того, что случайная величина примет эти значения.

ξ	X_1	X_2	$\dots$	X_i	$\dots$	X_n
P	p_1	p_2	$\dots$	p_i	$\dots$	p_n

еще одно определение дискретной случайной величины:

пусть задано дискретное вероятностное пространство (множество элементарных исходов дискретно): $\Omega = \{\omega_i\}_{i \in I}$

$P(x_i) = \sum\limits_{j \in I} P(\omega_j : \xi(\omega_j) = x_i) \neq 0$ - несколько событий могут отобразиться в один x_i

Тогда СВ **дискретна** и задаётся парами $\{x_i, P_i\}_{i \in I}$

Пример задания ДСВ:

ξ	0	1
P	0.5	0.5

17 функция распределения дискретной случайной величины и ее свойства

ℵ функция распределения случайной величины ξ — такая функция $F(x)$, значение которой в точке x равно вероятности события $\{\xi < x\}$, т. е. события, состоящего только из тех элементарных исходов ω, для которых $\xi(\omega) < x$:

$$F(x) = P(\xi < x)$$

свойства:

1. $0 \le F(x) \le 1$ (т. к. функция распределения это вероятность).
2. $F(x)$ — неубывающая функция (т. е. $F(x_1) \le F(x_2)$ при $x_1 < x_2$).
3. $F(-\infty) = 0, \quad F(+\infty) = 1$ (т. к. $\{\xi < -\infty\} = \varnothing$ и $\{\xi < +\infty\} = \Omega$).
4. $P(x_1 \le \xi < x_2) = F(x_2) - F(x_1)$ (т. к. событие $\{\xi < x_2\}$ при $x_1 < x_2$ представляет собой объединение двух непересекающихся событий: $\{\xi < x_1\}$ — случайная величина приняла значение, меньшее x_1, и $\{x_1 \le \xi < x_2\}$ — случайная величина приняла значение в интервале $[x_1, x_2)$. по аксиоме сложения получается, что сумма вероятностей этих "мини-событий" равна вероятности "большого" события).
5. $P(\xi \ge x) = 1 - F_\xi(x)$ (потому что это вероятности, че-то дополнение до события и так далее).
6. $F(x) = F(x - 0)$, т. е. $F(x)$ — непрерывная слева.
   док-во: пусть $x_1, \ \dots, \ x_n, \ \dots$ — возрастающая последовательность чисел, сходящаяся к x. тогда событие $A = \{\xi < x\}$ является счетным объединением несовместных событий

   $$A_1 = \{\xi < x_1\}, \quad A_n = \{x_{n-1} \le \xi < x_n\} \quad (n = 2, \ 3, \ \dots),$$

   то есть $A = A_1 + \ \dots \ + A_n + \ \dots$
   по аксиоме сложения $P(A) = P(A_1) + P(A_2) + \ \dots \ + P(A_n) + \ \dots$
   $= F(x_1) + (F(x_2) - F(x_1)) + \ \dots \ + (F(x_n) - F(x_{n-1})) + \ \dots = \lim\limits_{n \to \infty} F(x_n).$
   следовательно, $F(x)$ — непрерывная слева.
   (я тоже ничего не поняла.)

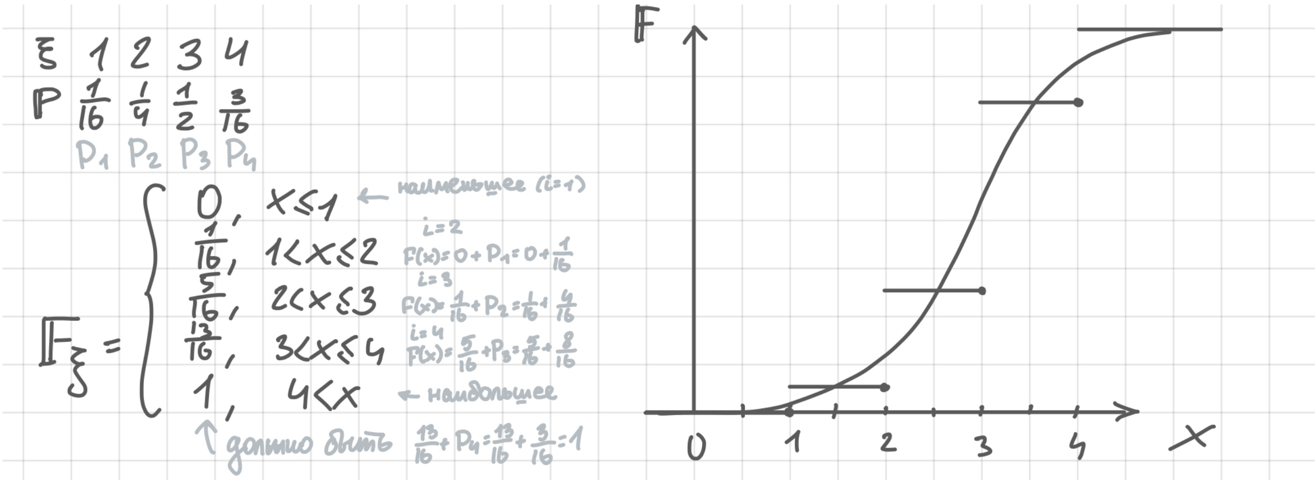
замечания

любая функция, которая удовлетворяет этим свойствам — $F_\xi(x)$.

как посчитать

всегда помните, что в формуле знак строго меньше $<$, а не меньше или равно $\leq$!!! это значит, что в кусочно-заданной $F(x)$ везде будет $\leq$ и $>$. это не супер логично, но если вы не понимаете логику, то просто заучите.

1. берем табличку случайной величины с n значениями $x_1, \ldots, x_n$ и вероятностями $P_1, \ldots, P_n$.
2. смотрим на самое маленькое значение, пусть оно равно x_1. $F_\xi(x) = 0$ при $x \leq x_1$.
3. смотрим на самое большое значение, пусть оно равно x_n. тогда $F_\xi(x) = 1$ при $x > x_n$.
4. для всех остальных промежутков $x_{i-1} < x \leq x_i$ нужно очень-очень аккуратно подумать мозгом. $F(x) = P(\xi < x)$. значит, так как вот тут у нас $<$, значит в функции надо рассматривать x в промежутке $(x_{i-1}, x_i]$. получается, что $F(x) = (\text{предыдущее } F(x)) + P_{i-1}$.



18 равномерное распределение на множестве чисел

$\aleph$ случайная величина ξ имеет равномерное распределение на множестве $\{1, \ldots, n\}$, если вероятность всех значений равна $\frac{1}{n}$.

$$P(\xi = k) = \frac{1}{n}, \quad \text{где } k = 1, \ldots, n$$

замечание:

вот это равномерное распределение относится только для дискретной случайной величины, так как непрерывная определена не на множестве чисел.

19 биномиальное распределение

$\aleph$ дискретная случайная величина ξ распределена по биномиальному закону, если она принимает значения $0,\ 1,\ \ldots,\ n$ в соответствии со следующим правилом:
$0 < p < 1, \quad k \in \{0,\ 1,\ \ldots,\ n\}$

$$P(\xi = k) = C_n^k p^k (1-p)^{n-k}$$

биномиальное распределение — распределение числа успехов в испытаниях Бернулли с вероятностью успеха p и неудачи $1-p$.

математическое ожидание для биномиального распределения:

$$M_\xi = np$$

20 распределение Пуассона

ℵ дискретная случайная величина ξ распределена по закону Пуассона с параметром пуассоновского распределения $\lambda > 0$, если она принимает целые неотрицательные значения по следующему правилу:
$\xi \in \{0,\ 1,\ \dots,\ n\}$

$$P(\xi = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$$

доказательство того, что это выражение является законом распределения. для этого нужно доказать, что сумма вероятностей случайных величин равна 1 (т. е. что $\sum P = 1$).
$P_k \geq 0$, т. к. $k! > 0$, $\lambda^k > 0$ и $e^{-\lambda} > 0$.

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{=e^{\lambda}} = e^0 = 1$$

$\sum P = 1 \Rightarrow$ подходит для закона распределения. ура!

примечание для тех, кто не понял, почему оно равно e^{λ}:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

это разложение в ряд Тейлора.

замечание

если в какой-то задаче $n \cdot p \cdot q \leq 9$, то следует применить закон распределения Пуассона.

21 теорема Пуассона

пусть есть последовательность опытов Бернулли, где p_n — вероятность успеха, а случайная величина ξ определяется количеством успехов. тогда если $\lim\limits_{n\to\infty} n \cdot p_n = \lambda > 0$, то

$$\lim_{n\to\infty} P(\omega : \xi(\omega) = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

то есть биномиальное распределение можно заменить на пуассоновское.

доказательство

$n \cdot p_n = \lambda_n \Rightarrow p_n = \dfrac{\lambda_n}{n}$.

по формуле Бернулли:

$$\lim_{n\to\infty} P(\xi = k) = C_n^k \cdot (p_n)^k \cdot (1 - p_n)^{n-k} = \frac{n!}{k!(n-k)!} \cdot \left(\frac{\lambda_n}{n}\right)^k \cdot \left(1 - \frac{\lambda_n}{n}\right)^{n-k} =$$
$$\frac{1}{k!} \cdot \frac{(n-k+1) \cdot (n-k+2) \cdot \ \ldots \ \cdot n}{n^k} \cdot (\lambda_n)^k \cdot \left(1 - \frac{\lambda_n}{n}\right)^{n-k} =$$

1. $\lim\limits_{n\to\infty} \dfrac{(n-k+1) \cdot (n-k+2) \cdot \ \ldots \ \cdot n}{n^k} = \left(\lim\limits_{n\to\infty} \dfrac{n-k+1}{n}\right) \cdot \left(\lim\limits_{n\to\infty} \dfrac{n-k+2}{n}\right) \cdot \ \ldots \ \cdot \left(\lim\limits_{n\to\infty} \dfrac{n}{n}\right) = 1$
2. $\lim\limits_{n\to\infty} \left(1 - \dfrac{\lambda}{n}\right)^{n-k} = e^{-\lambda}$

$$= \frac{\lambda_n^k}{k!} e^{-\lambda}.$$

и все собственно.

1 многомерное распределение, дискретный случай

пусть на одном и том же вероятностном пространстве $(\Omega, \mathfrak{A}, P)$ задано n случайных величин: $\xi_1 = \xi_1(\omega), \ldots, \xi_n = \xi_n(\omega)$.

$\aleph$ упорядоченный набор случайных величин $(\xi_1, \ldots, \xi_n)$ называется многомерной (n-мерной) случайной величиной или случайным вектором.
ξ_i $(i \in I)$ — компонента случайного вектора.

функция распределения:
для одномерной случайной величины $F_\xi(x) = P(\xi < x)$
для многомерной: $F_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n) = P\underbrace{(\xi_1 < x_1, \ldots, \xi_n < x_n)}_{\in \mathfrak{A}}$ — функция распределения n-мерного случайного вектора (также называют совместной функцией распределения случайных величин $\xi_1, \ldots, \xi_n$).

$\aleph$ n-мерная случайная величина распределена по дискретному закону, если $(X_i, P_i)_{i \in I}$, где $X_i = (x_i^1, \ldots, x_i^n)$, $P_i = P(\xi = X_i)$ — закон распределения дискретной случайной величины (т. е. все ξ_i — дискретные случайные величины.)

для двумерной случайной величины:
$i = 1, \ldots, n$, $j = 1, \ldots, m$

$$\sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} = 1$$

2 независимость случайной величины

из конспекта

ℵ случайные величины ξ и η попарно независимы, если $F_{\xi\eta}(x,\ y) = F_{\xi}(x) \cdot F_{\eta}(y)$, т. е. $P(\xi = x_i,\ \eta = y_i) = P(\xi = x_i) \cdot P(\eta = y_i)$ для всех $x_i,\ y_i$.

ℵ случайные величины $\xi_1,\ \dots,\ \xi_n$ независимы в совокупности, если $F_{\xi_1, \dots, \xi_n}(x_1,\ \dots,\ x_n) = F_{\xi_1}(x_1) \cdot\ \dots\ \cdot F_{\xi_n}(x_n)$ для всех $x_1,\ \dots,\ x_n$.

еще одно определение для двумерной случайной величины:
$M = (\Omega,\ \mathfrak{A},\ P)$, на $\mathbb{R}^1$ определены ξ и η;
$A = \{\omega : \xi(\omega) < x\}$
$B = \{\omega : \eta(\omega) < y\}$
если A и B независимы для любых $x,\ y$, то ξ и η независимы.

из учебника

ℵ назовем случайные величины ξ и η независимыми, если совместная функция распределения $F_{\xi,\eta}(x,\ y)$ представляется в виде произведения одномерных функций распределения $F_{\xi}(x)$ и $F_{\eta}(y)$:

$$F_{\xi,\eta}(x,\ y) = F_{\xi}(x)F_{\eta}(y)$$

понятие независимости случайных величин представляет собой перенос понятия независимости событий на случайные величины и отражает отсутствие связи между случайными величинами ξ и η, хотя они и заданы на одном и том же вероятностном пространстве $(\Omega,\ \mathfrak{A},\ P)$ и определяют двумерный случайный вектор. то есть получив значение η, мы не можем ничего сказать о значении ξ.

ℵ случайные величины $\xi_1,\ \dots,\ \xi_n$ называются независимыми в совокупности, если:

$$F_{\xi_1, \dots, \xi_n}(x_1,\ \dots,\ x_n) = F_{\xi_1}(x_1) \cdot\ \dots\ \cdot F_{\xi_n}(x_n)$$

так же, как и для событий, из независимости в совокупности следует попарная независимость, но не наоборот. здесь даже можно привести такой же пример, как в 4 билетике из первой части, превратив вероятности выпадения сторон тетраэдра в случайные величины.

Разумеется, так же, как и для событий, из попарной независимости не следует независимость случайных величин в совокупности.

Пример 17. Свяжем с бросанием правильного раскрашенного тетраэдра (пример 11 в гл. 3) три случайные величины: ξ_1, ξ_2 и ξ_3, каждая из которых может принимать значения 0 или 1, причем $\xi_1 = 1$, если тетраэдр упал на грань, в раскраске которой присутствует синий цвет, и $\xi_1 = 0$ в противном случае. Аналогично, ξ_2 характеризует наличие красного цвета, а ξ_3 — зеленого. Нетрудно видеть, что случайные величины ξ_1, ξ_2 и ξ_3 будут попарно независимы, но зависимы в совокупности. □

вот вам еще куча всякой фигни:

Для проверки независимости компонент многомерных дискретных и непрерывных случайных векторов обычно бывают удобными другие эквивалентные определения независимости (доказательство эквивалентности приводимых ниже определений предоставляем читателю). Так, *дискретные случайные величины ξ и η независимы, если для всех возможных значений X_i и Y_j*

$$p_{ij} = \mathsf{P}\{\xi = X_i, \eta = Y_j\} = \mathsf{P}\{\xi = X_i\}\mathsf{P}\{\eta = Y_j\} = p_{\xi i} p_{\eta j}.$$

Таким образом, числа успехов μ_1 и μ_2 в первом и втором испытаниях представляют собой независимые случайные величины. Впрочем, иного и нельзя было ожидать из самого определения схемы Бернулли. Читатель может самостоятельно убедиться в том, что независимы в совокупности случайные величины $\mu_1, \ldots, \mu_n$ — числа успехов в первом, ..., n-м испытаниях Бернулли. □

Непрерывные случайные величины ξ и η независимы, если для всех x и y

$$p_{\xi,\eta}(x, y) = p_\xi(x)\, p_\eta(y).$$

Отметим здесь же, что если независимые случайные величины ξ и η непрерывны, то и двумерная случайная величина (ξ, η) обязана быть непрерывной (ср. с примером 10).

3 распределение суммы независимых дискретных случайных величин

ℵ сумма случайных величин ξ и η — случайная величина, которая принимает все возможные значения вида $x_i + y_i$ (где $i \in I$ — дискретное) с вероятностью P_{ij} $(j \in I)$ того, что $\xi = x_i,\ \eta = y_i$.

$$P_{ij} = P(\xi = x_i,\ \eta = y_i)$$

пример

даны законы распределения:

ξ	0	1	2
P_ξ	0.5	0.3	0.2

η	0	1	2
P_η	0.4	0.1	0.3

тогда $F_{\xi,\,\eta}$:

$\xi+\eta$	$0+0$	$0+1$	$0+2$	$\dots$
$F_{\xi+\eta}$	0.2	0.05	0.15	$\dots$

```
но это типа... хз, как-то странно и мало. я не уверена,
нужно уточнить на консультации.
```

4 математическое ожидание дискретной случайной величины

ℵ математическое ожидание (среднее значение) M_ξ дискретной случайной величины ξ — сумма произведений значений X_i случайной величины на вероятности $p_i = P(\xi = X_i)$, с которыми величина принимает эти значения:

$$M(\xi) = \sum_i X_i p_i$$

при этом, если случайная величина ξ принимает счетное число значений, то необходимо, чтобы

$$\sum_{i=1}^{\infty} |X_i| p_i < \infty$$

то есть ряд должен сходиться абсолютно. в ином случае говорят, что матожидания не существует.

математическое ожидание — это такое среднее значение случайной величины. в физике это центр тяжести какого-то объекта или набора объектов.

для биномиального распределения $M(\xi) = np$, для пуассоновского — $M(\xi) = \lambda$.

пример, когда матожидание не существует:

Пример 5. Положительная целочисленная случайная величина ξ имеет закон распределения $p_i = \mathsf{P}\{\xi = i\} = 1/[i(i+1)]$ $(i = 1, 2, \ldots)$. Тогда

$$\sum_{i=1}^{\infty} i p_i = \sum_{i=1}^{\infty} \frac{i}{i(i+1)} = \sum_{i=1}^{\infty} \frac{1}{i+1} = \infty$$

и, значит, математическое ожидание случайной величины ξ не существует. □

свойства

1. $M(c) = c \cdot 1 = c$
   если случайная величина ξ принимает всего одно значение c с вероятностью, равной единице, т. е. по сути дела не является случайной величиной, то его матожидание равно c. супер логично.

2. $\xi \geq 0 \Rightarrow M(\xi) \geq 0$
   тоже супер логично, не может же среднее значение положительной случайной величины быть отрицательным?

3. $\xi \geq \eta \Rightarrow M(\xi) \geq M(\eta)$

4. $M(a \cdot \xi) = a \cdot M(\xi)$

5. $M(\xi + \eta) = M(\xi) + M(\eta)$

$$\mathsf{M}\eta = \mathsf{M}(\xi_1 + \xi_2) = \sum_{i,j}(X_i + Y_j)\,p_{ij} = \sum_{i,j} X_i\,p_{ij} + \sum_{i,j} Y_j\,p_{ij} =$$
$$= \sum_i X_i \sum_j p_{ij} + \sum_j Y_j \sum_i p_{ij} = \sum_i X_i\,p_{\xi_1 i} + \sum_j Y_j\,p_{\xi_2 j}$$

6. $M(a\xi + b\eta) = aM\xi + bM(\eta)$

7. если ξ и η независимы, то $M(\xi \cdot \eta) = M(\xi) \cdot M(\eta)$

8. $\eta = f(\xi), \quad M(\eta) = \sum\limits_{i \in I} f(x_i)p_i$

5 индикатор случайной величины

ℵ индикатором случайного события A называется случайная величина $IA(\omega)$, показывающая, наступило ли событие ω.

$$I(A) = IA(\omega) = \begin{cases} 1, \ \omega \in A \\ 0, \ \omega \notin A \end{cases}$$

$I(A)$ — дискретная случайная величина, так как принимает значения 1 или 0.

свойства индикаторов

1. $M(IA) = P(A)$
2. $I(\varnothing) = 0$
3. $I(\Omega) = 1$
4. $I(AB) = I(A) \cdot I(B)$
5. $I(A + B) = I(A) + I(B) - I(AB)$
6. $I^n(A) = I(A)$, где $n \geq 0$

че это за фигня зачем это нужно??

пример

стрелок стреляет в мишень. событие A — стрелок попал в мишень; событие B — не попал. пусть стрелок выстрелил один раз и промахнулся. тогда $I(A) = 0, \ I(B) = 1.$

6 математическое ожидание для функции от случайной величины

пусть $\eta = f(\xi)$. обозначим $M(\eta) = M(f(\xi))$.
рассмотрим дискретную случайную величину ξ, принимающую значения $(x_1, \ \dots, \ x_n)$. тогда η принимает значения $(f(x_1), \ \dots, \ f(x_n))$ с теми же вероятностями $p_i = P(\xi = x_i)$.

$$M(\eta) = \sum_{i=1}^{n} f(x_i) \cdot p_i$$

замечание: матожидание существует только тогда, когда этот ряд сходится абсолютно.

аналогично для непрерывной случайной величины ξ, имеющей плотность $p_\xi(x)$.
$\eta = f(\xi)$

$$M(\eta) = \int\limits_{-\infty}^{+\infty} f(x) P_\xi(x) dx$$

замечание: интеграл должен сходиться абсолютно.

7 дисперсия дискретной случайной величины

ℵ дисперсия — числовая характеристика случайной величины, показывающая «разброс» этой величины вокруг ее среднего значения.

здесь, несмотря на порядок вопросов в билетах, нужно понятие второго (начального) момента дискретной случайной величины:

$$\alpha_2 = M(\xi^2) = \sum_i X_i^2 p_i$$

дисперсия дискретной случайной величины определяется по формуле:

$$D(\xi) = M(\xi - M(\xi))^2 = \sum_i (X_i - M(\xi))^2 p_i$$

свойства

1. если $\xi = c\ (const)$, то $D(\xi) = D(c) = 0$
2. $D(\xi) \geq 0$
3. $D(\xi) = M(\xi^2) - (M(\xi))^2$
   $D(\xi) = M(\xi - M(\xi))^2 = M(\xi^2 - 2\xi M(\xi) + (M(\xi))^2) =$
   $M(\xi^2) - 2(M(\xi))^2 + (M(\xi))^2 = M(\xi^2) - (M(\xi))^2$
4. $D(a\xi + b) = a^2 D(\xi)$
   $D(a\xi + b) = M(a\xi + b - M(a\xi + b))^2 = M(a\xi + b - aM(\xi) - b)^2 =$
   $M(a(\xi - M\xi))^2 = M(a^2(\xi - M\xi)^2) = a^2 M(\xi - M(\xi))^2 = a^2 D(\xi)$
5. если ξ и η — независимые, то $D(\xi + \eta) = D(\xi) + D(\eta)$
6. $D(\xi + \eta) = D(\xi) + D(\eta) + 2cov(\xi,\ \eta)$, где
   $cov(\xi,\ \eta) = M((\xi - M(\xi))(\eta - M(\eta)))$
7. если $\xi_1,\ \ldots,\ \xi_n$ — независимые (попарно, нет необходимости в совокупности), то

$$D(\sum_{i=1}^{n} \xi_i) = \sum_{i=1}^{n} D(\xi_i)$$

8. в общем случае:

$$D(\sum_{i=1}^{n} \xi_i) = \sum_{i=1}^{n} D(\xi_i) + 2 \sum_{1 \le i < j \le n} cov(\xi_i,\ \xi_j)$$

еще у нас есть такая штука, как среднее квадратичное отклонение $\sigma = \sqrt{D(\xi)}$. его еще называют среднеквадратическое отклонение, стандартное отклонение.
я хз, будет ли оно дальше как отдельная тема, поэтому вот, держите.

Заметим, что дисперсия $\mathsf{D}\xi$ имеет размерность квадрата размерности случайной величины ξ. Для практических же целей удобно иметь меру разброса, размерность которой совпадает с размерностью ξ. В качестве такой меры естественно использовать $\sigma = \sqrt{\mathsf{D}\xi}$, которую называют *средним квадратичным отклонением* случайной величины ξ (иногда также *стандартом* или *стандартным отклонением*).

8 ковариация дискретной случайной величины

ковариация — мера зависимости двух случайных величин друг от друга.

$$cov(\xi_1,\ \xi_2) = M((\xi_1 - M(\xi_1))(\xi_2 - M(\xi_2)))$$

для дискретных случайных величин:

$$cov(\xi_1,\ \xi_2) = \sum_{i,\ j}(X_i - M(\xi_1))(Y_j - M(\xi_2))p_{ij}$$

свойства

1. $cov(\xi,\ \xi) = D(\xi)$
2. если ξ_1 и ξ_2 независимы, то $cov(\xi_1,\ \xi_2) = 0$. обратное утверждение неверно.
3. $cov(a_1\xi_1 + b_1,\ a_2\xi_2 + b_2) = a_1 a_2 cov(\xi_1,\ \xi_2)$
4. $-\sqrt{D(\xi_1)D(\xi_2)} \le cov(\xi_1,\ \xi_2) \le \sqrt{D(\xi_1)D(\xi_2)}$, причём крайние значения ковариация принимает только тогда, когда случайные величины линейно зависимы.
5. $cov(\xi_1,\ \xi_2) = M(\xi_1\xi_2) - M(\xi_1)M(\xi_2)$

9 корреляция дискретной случайной величины

если ковариация двух случайных величин равна нулю, то эти величины называются некоррелированными.

ℵ коэффициентом корреляции случайных величин называется число $\rho - \rho(\xi,\ \eta)$, определяемое выражением:

$$\rho = \frac{cov(\xi,\ \eta)}{\sqrt{D(\xi)D(\eta)}}$$

свойства:

1. $\rho(\xi,\ \xi) = 1$
2. если ξ и η независимы, то $\rho(\xi,\ \eta) = 0$
3. $\rho(a_1\xi_1 + b_1,\ a_2\xi_2 + b_2) = \dfrac{a_1a_2cov(\xi_1,\ \xi_2)}{\sqrt{a_1^2D(\xi_1) \cdot a_2^2D(\xi_2))}} = \pm p(\xi_1,\ \xi_2)$
4. $-1 \leq \rho(\xi,\ \eta) \leq 1$, причем принимает крайние значения только тогда, когда случайные величины линейно зависимы.

10 дисперсия суммы дискретных случайных величин

1. если ξ и η — независимые, то $D(\xi+\eta)=D(\xi)+D(\eta)$
2. $D(\xi+\eta)=D(\xi)+D(\eta)+2cov(\xi,\ \eta)$, где $cov(\xi,\ \eta)=M((\xi-M(\xi))(\eta-M(\eta)))$
3. если $\xi_1,\ \ldots,\ \xi_n$ — независимые (попарно, нет необходимости в совокупности), то

$$D(\sum_{i=1}^{n}\xi_i)=\sum_{i=1}^{n}D(\xi_i)$$

4. в общем случае:

$$D(\sum_{i=1}^{n}\xi_i)=\sum_{i=1}^{n}D(\xi_i)+2\sum_{1\le i<j\le n}cov(\xi_i,\ \xi_j)$$

11 моменты дискретной случайной величины

здесь все формулы для **дискретной** случайной величины. для непрерывной другие формулы.
еще примечание: можно разными буквами все это обозначать. в учебнике это m и $\mathring{m}$, в конспектах α и μ.

как уже рассказывалось в 7 вопросике,
$\aleph$ вторым начальным моментом α_2 случайной величины ξ называется математическое ожидание квадрата ξ:

$$\alpha_2 = M(\xi^2) = \sum_i X_i^2 p_i$$

$\aleph$ моментом α_k порядка k (k-м моментом) называется математическое ожидание k-й степени случайной величины ξ:

$$\alpha_k = M(\xi^k) = \sum_i X_i^k p_i$$

$\aleph$ центральным моментом α_k порядка k (k-м центральным моментом) называется математическое ожидание k-й степени случайной величины $\xi - M(\xi)$:

$$\mu_k = M(\xi - M(\xi))^k = \sum_i (X_i - M(\xi))^k p_i$$

частные случаи

1. $\alpha_1 = M(\xi)$
2. $\mu_1 = M(\xi - M(\xi)) = 0$
3. $\mu_2 = M(\xi - M(\xi))^2 = D(\xi) = \alpha_2 - (\alpha_1)^2$

также есть абсолютные моменты, где нужно подставлять случайную величину по модулю:

- $M(|\xi|^k)$ — k-й абсолютный начальный момент;
- $M(|\xi - M(\xi)|^k)$ — k-й абсолютный центральный момент.

12 случайная величина непрерывного типа

$\aleph$ непрерывная случайная величина (распределенная по непрерывному типу) — случайная величина ξ, функцию распределения $F(\xi)$ которой можно представить в виде

$$F(x) = \int_{-\infty}^{x} p(\tau)\, d\tau$$

⋆ τ это не т, это греческая тау.

$\aleph$ функция $p(\tau)$ называется плотностью распределения (вероятностей) случайной величины ξ. еще ее называют дифференциальной функцией распределения; мы часто будем называть ее просто плотностью, но это уже наша вольность, а не реальный термин.

соответственно,

$$p_\xi(x) = F'_\xi(x)$$

свойства плотности распределения

1. $p(x) \geq 0$
   функция распределения неубывающая, поэтому ее производная не может быть отрицательной.
2. $\int\limits_{-\infty}^{\infty} p(x)\, dx = 1$
   то же самое, что $F_\xi(\infty) = 1$
3. $P(x_1 \leq \xi \leq x_2) = \int\limits_{x_1}^{x_2} p(x)\, dx$

13 функция распределения случайной величины

```
разве не та же фигня, что 17 вопрос 1 части?...
```

ℵ функция распределения случайной величины ξ — такая функция $F(x)$, значение которой в точке x равно вероятности события $\{\xi < x\}$, т. е. события, состоящего только из тех элементарных исходов ω, для которых $\xi(\omega) < x$:

$$F_\xi(x) = P(\omega : \xi(\omega) < x) = P(\xi < x)$$

$(\Omega,\ \mathfrak{A},\ P)$ — вероятностное пространство; $\omega \in \mathfrak{A}$.

$$F_\xi(x) = \begin{cases} \sum\limits_{i \in I} P(\xi < x_i), & \text{если } \xi \text{ дискретная} \\ \int\limits_{-\infty}^{x} p(\tau)\, d(\tau), & \text{если } \xi \text{ непрерывная} \end{cases}$$

14 свойства функции непрерывной случайной величины

$$F(x) = \int\limits_{-\infty}^{x} p(\tau)\, d\tau$$

свойства:

1. $0 \le F(x) \le 1$ (т. к. функция распределения это вероятность).
2. $F(x)$ — неубывающая функция (т. е. $F(x_1) \le F(x_2)$ при $x_1 < x_2$).
3. $F(-\infty) = 0, \quad F(+\infty) = 1$ (т. к. $\{\xi < -\infty\} = \varnothing$ и $\{\xi < +\infty\} = \Omega$).
4. $P(x_1 \le \xi < x_2) = F(x_2) - F(x_1)$ (т. к. событие $\{\xi < x_2\}$ при $x_1 < x_2$ представляет собой объединение двух непересекающихся событий: $\{\xi < x_1\}$ — случайная величина приняла значение, меньшее x_1, и $\{x_1 \le \xi < x_2\}$ — случайная величина приняла значение в интервале $[x_1, x_2)$. по аксиоме сложения получается, что сумма вероятностей этих "мини-событий" равна вероятности "большого" события).
5. $P(\xi \ge x) = 1 - F_\xi(x)$ (потому что это вероятности, че-то дополнение до события и так далее).
6. $F(x) = F(x-0)$, т. е. $F(x)$ — непрерывная слева.
   док-во: пусть $x_1, \, \ldots, \, x_n, \, \ldots$ — возрастающая последовательность чисел, сходящаяся к x. тогда событие $A = \{\xi < x\}$ является счетным объединением несовместных событий

   $$A_1 = \{\xi < x_1\}, \quad A_n = \{x_{n-1} \le \xi < x_n\} \quad (n = 2, \, 3, \, \ldots),$$

   то есть $A = A_1 + \, \ldots \, + A_n + \, \ldots$
   по аксиоме сложения $P(A) = P(A_1) + P(A_2) + \, \ldots \, + P(A_n) + \, \ldots$
   $= F(x_1) + (F(x_2) - F(x_1)) + \, \ldots \, + (F(x_n) - F(x_{n-1})) + \, \ldots = \lim\limits_{n \to \infty} F(x_n).$
   следовательно, $F(x)$ — непрерывная слева.
   (я тоже ничего не поняла.)

замечания

любая функция, которая удовлетворяет этим свойствам — $F_\xi(x)$.

15 математическое ожидание непрерывной случайной величины и ее свойства

ℵ математическое ожидание (среднее значение) M_ξ непрерывной случайной величины ξ — это интеграл, вычисляемый следующим образом: (здесь $p(x)$ — плотность распределения)

$$M(\xi) = \int_{-\infty}^{\infty} x\, p(x)\, dx$$

при этом для существования математического ожидания необходимо, чтобы

$$\int_{-\infty}^{\infty} |x|\, p(x)\, dx < \infty$$

то есть ряд должен сходиться абсолютно. в ином случае говорят, что матожидания не существует.

математическое ожидание — это такое среднее значение случайной величины. в физике это центр тяжести стержня, который в какой-то точке x имеет плотность $p(x)$.

свойства

1. $M(c) = c \cdot 1 = c$
   если случайная величина ξ принимает всего одно значение c с вероятностью, равной единице, т. е. по сути дела не является случайной величиной, то его матожидание равно c. супер логично.

2. $\xi \geq 0 \Rightarrow M(\xi) \geq 0$
   тоже супер логично, не может же среднее значение положительной случайной величины быть отрицательным?

3. $\xi \geq \eta \Rightarrow M(\xi) \geq M(\eta)$

4. $M(a\xi + b) = aM(\xi) + b$

Далее, найдем математическое ожидание случайной величины $\eta = a\xi + b$ $\quad (g(x) = ax + b)$. Рассматривая, например, непрерывный случай, имеем

$$\mathsf{M}\eta = \mathsf{M}(a\xi + b) = \int_{-\infty}^{\infty} (ax + b)\, p_\xi(x)\, dx = a\int_{-\infty}^{\infty} x\, p_\xi(x)\, dx + b\int_{-\infty}^{\infty} p_\xi(x)\, dx,$$

т. е.

2. $\mathsf{M}(a\xi + b) = a\mathsf{M}\xi + b.$

5. $M(\xi + \eta) = M(\xi) + M(\eta)$

6. $M(a\xi + b\eta) = aM\xi + bM(\eta)$

7. если ξ и η независимы, то $M(\xi \cdot \eta) = M(\xi) \cdot M(\eta)$

8. $\eta = f(\xi), \quad M(\eta) = \int_{-\infty}^{\infty} f(x)\, p_\xi(x)\, dx$

пример

Пример 6. Найдем математическое ожидание равномерно распределенной на отрезке $[a, b]$ случайной величины ξ. Поскольку в этом случае $p(x) = 0$ при $x < a$ и $x > b$, то

$$\mathsf{M}\xi = \int_{-\infty}^{\infty} x\, p(x)\, dx = \int_a^b \frac{x}{b - a}\, dx = \frac{1}{b - a} \cdot \frac{1}{2}\,(b^2 - a^2) = \frac{b + a}{2}\,.$$

Как и следовало ожидать, $\mathsf{M}\xi$ совпадает с центром отрезка $[a, b]$. $\square$

16 дисперсия непрерывной случайной величины и ее свойства

ℵ дисперсия — числовая характеристика случайной величины, показывающая «разброс» этой величины вокруг ее среднего значения.

здесь нужно понятие второго (начального) момента непрерывной случайной величины:

$$\alpha_2 = M(\xi^2) = \int\limits_{-\infty}^{\infty} x^2 \, p(x) \, dx$$

дисперсия непрерывной случайной величины определяется по формуле:

$$D(\xi) = M(\xi - M(\xi))^2 = \int\limits_{-\infty}^{\infty} (x - M(\xi))^2 \, p(x) \, dx$$

свойства

1. если $\xi = c \ (const)$, то $D(\xi) = D(c) = 0$
2. $D(\xi) \geq 0$
3. $D(\xi) = M(\xi^2) - (M(\xi))^2$
   $D(\xi) = M(\xi - M(\xi))^2 = M(\xi^2 - 2\xi M(\xi) + (M(\xi))^2) =$
   $M(\xi^2) - 2(M(\xi))^2 + (M(\xi))^2 = M(\xi^2) - (M(\xi))^2$
4. $D(a\xi + b) = a^2 D(\xi)$
   $D(a\xi + b) = M(a\xi + b - M(a\xi + b))^2 = M(a\xi + b - aM(\xi) - b)^2 =$
   $M(a(\xi - M\xi))^2 = M(a^2(\xi - M\xi)^2) = a^2 M(\xi - M(\xi))^2 = a^2 D(\xi)$
5. если ξ и η — независимые, то $D(\xi + \eta) = D(\xi) + D(\eta)$
6. $D(\xi + \eta) = D(\xi) + D(\eta) + 2cov(\xi, \ \eta)$, где
   $cov(\xi, \ \eta) = M((\xi - M(\xi))(\eta - M(\eta)))$
7. если $\xi_1, \ \ldots, \ \xi_n$ — независимые (попарно, нет необходимости в совокупности), то

$$D(\sum_{i=1}^{n} \xi_i) = \sum_{i=1}^{n} D(\xi_i)$$

8. в общем случае:

$$D(\sum_{i=1}^{n} \xi_i) = \sum_{i=1}^{n} D(\xi_i) + 2 \sum_{1 \le i < j \le n} cov(\xi_i,\ \xi_j)$$

еще у нас есть такая штука, как среднее квадратичное отклонение $\sigma = \sqrt{D(\xi)}$. его еще называют среднеквадратическое отклонение, стандартное отклонение.
я хз, будет ли оно дальше как отдельная тема, поэтому вот, держите.

Заметим, что дисперсия $\mathsf{D}\xi$ имеет размерность квадрата размерности случайной величины ξ. Для практических же целей удобно иметь меру разброса, размерность которой совпадает с размерностью ξ. В качестве такой меры естественно использовать $\sigma = \sqrt{\mathsf{D}\xi}$, которую называют *средним квадратичным отклонением* случайной величины ξ (иногда также *стандартом* или *стандартным отклонением*).

17 начальные и центральные моменты непрерывных случайных величин

здесь все формулы для **непрерывной** случайной величины. для дискретной другие формулы.
еще примечание: можно разными буквами все это обозначать. в учебнике это m и $\mathring{m}$, в конспектах α и μ.

как уже рассказывалось в 16 вопросике,
ℵ вторым начальным моментом α_2 случайной величины ξ называется математическое ожидание квадрата ξ:

$$\alpha_2 = M(\xi^2) = \int\limits_{-\infty}^{\infty} x^2\, p(x)\, dx$$

ℵ моментом α_k порядка k (k-м моментом) называется математическое ожидание k-й степени случайной величины ξ:

$$\alpha_k = M(\xi^k) = \int\limits_{-\infty}^{\infty} x^k\, p(x)\, dx$$

ℵ центральным моментом α_k порядка k (k-м центральным моментом) называется математическое ожидание k-й степени случайной величины $\xi - M(\xi)$:

$$\mu_k = M(\xi - M(\xi))^k = \int\limits_{-\infty}^{\infty} (x - M(\xi))^k\, p(x)\, dx$$

частные случаи

1. $\alpha_1 = M(\xi)$
2. $\mu_1 = M(\xi - M(\xi)) = 0$
3. $\mu_2 = M(\xi - M(\xi))^2 = D(\xi) = \alpha_2 - (\alpha_1)^2$

также есть абсолютные моменты, где нужно подставлять случайную величину по модулю:

- $M(|\xi|^k)$ — k-й абсолютный начальный момент;
- $M(|\xi - M(\xi)|^k)$ — k-й абсолютный центральный момент.

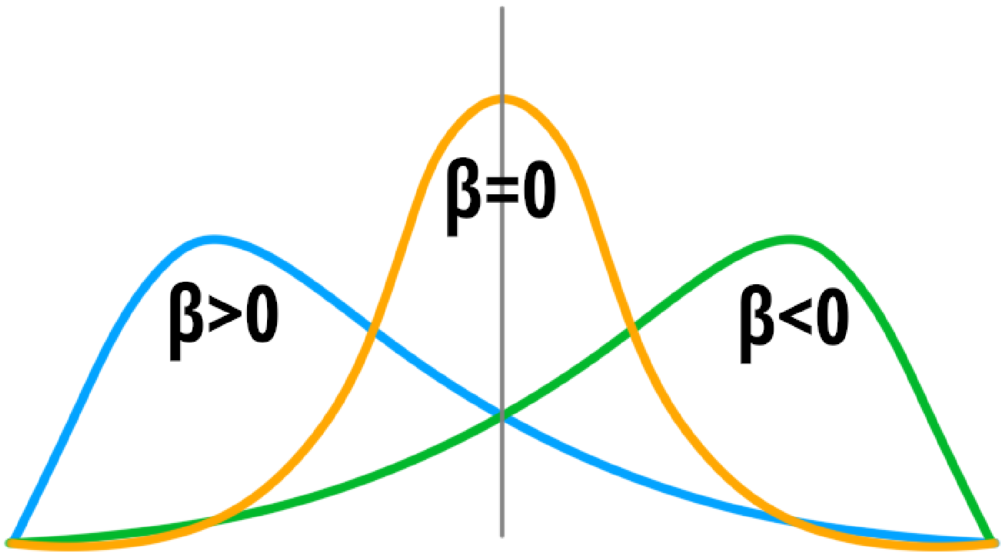
18 коэффициенты асимметрии и эксцесса

```
запомните, как пишется. асимметрия, эксцесс. "математик
должен быть грамотным." (с) Шмидт
```

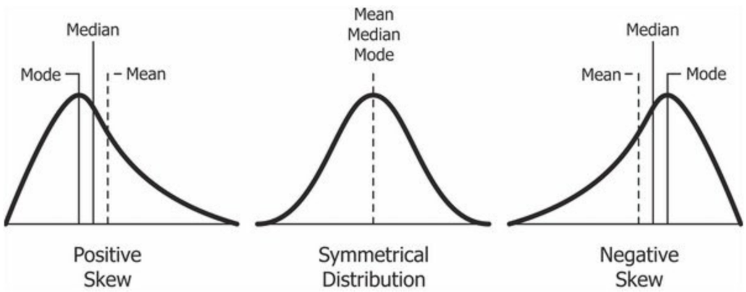
ℵ коэффициент асимметрии показывает, насколько распределение асимметрично относительно прямой $x = M(\xi)$.

$$\beta(\xi) = \frac{\mu_3}{\sigma^3}$$

где μ_3 — третий центральный момент ($\mu_3 = M(\xi - M(\xi))^3$), $\sigma = \sqrt{D(\xi)}$ — среднее квадратичное отклонение.



у всех трех распределений на картинке матожидание равно 0. если распределение симметрично относительно математического ожидания, то его коэффициент асимметрии равен нулю. также, когда $\beta(\xi) = 0$, матожидание, мода и медиана совпадают.

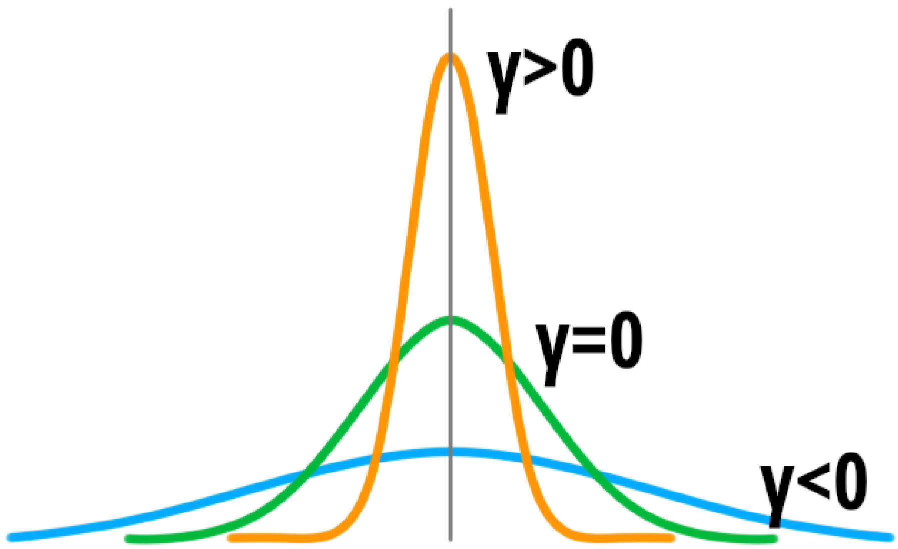


положительная асимметрия — когда правая часть графика длиннее, отрицательная асимметрия — когда левая часть графика длиннее.

ℵ коэффициент эксцесса — мера остроты пика распределения случайной величины.

$$\gamma = \frac{\mu_4}{\sigma^4} - 3$$

⋆ 3 отнимаем, потому что ~~три за экзамен будет~~ для нормального распределения коэффициент эксцесса равен 3, а мы любим с нормальным распределением все сравнивать.



дополнительное.

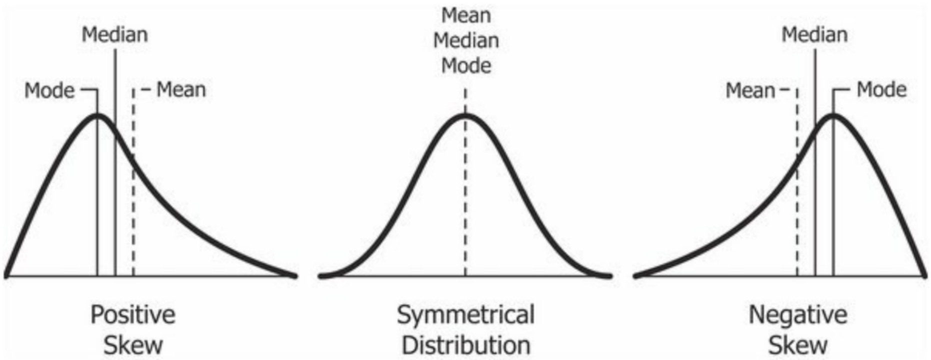
свойства:

1. $\gamma \in [-2,\ \infty)$
2. пусть $\xi_1,\ \ldots,\ \xi_n$ — независимые случайные величины с равной дисперсией. пусть $\eta = \sum\limits_{i=1}^{n} \xi_i$. тогда $\gamma_\eta = \frac{1}{n^2} \sum\limits_{i=1}^{n} \gamma_\xi$

   то есть если случайная величина является суммой независимых случайных величин с равной дисперсией, то ее коэффициент эксцесса равен сумме коэффициентов эксцесса слагаемых случайных величин.

19 мода и медиана случайной величины

mean — среднее значение
mode — мода
median — медиана



ℵ модой **непрерывной** случайной величины называется точка (локального) максимума плотности распределения $p(x)$.

ℵ модой **дискретной** случайной величины, значения $X_1, \ldots, X_n$ которой расположены по возрастанию, называется такое значение X_i, что $p_{i-1} < p_i$ и $p_{i+1} < p_i$.

в зависимости от количества мод распределения бывают:

- унимодальные (1 мода);
- бимодальные (2 моды);
- полимодальные (3 моды и более).

доп. материал:

ℵ наивероятнейшее значение — мода, доставляющая глобальный максимум вероятности (для дискретных случайных величин) или плотности распределения (для непрерывных случайных величин). если распределение унимодальное, то мода также будет наивероятнейшим значением.

в учебнике также отмечено, что мода и наивероятнейшее значение введены скорее для наглядности, чем для каких-либо практических целей.

ℵ медиана — такое значение l, что $P(\xi < l) = P(\xi \geq l) = \frac{1}{2}$.

медиана разделяет выборку на две части.
часто медиана бывает более характеризующей, чем среднее значение, потому что одно очень большое значение случайной величины (в случае с положительной асимметрией, например) может немного "портить" статистику. например, в лесу среди обычных таких деревьев может вдруг затесаться секвойя, и по среднему значению лес получится выше среднего, хотя медиана покажет, что он вполне обычный.

20 стандартизация и централизация случайной величины

```
очень, *очень* не уверена в этом билете. верьте наполовину.
```

ℵ центрированной случайной величиной называется отклонение случайной величины от ее математического ожидания:

$$\xi_0 = \xi - M(\xi)$$

центрирование случайной величины: $\eta = \xi - M(\xi)$. $M(\eta) = 0$.

ℵ стандартизованная случайная величина — это случайная величина, математическое ожидание которой равно нулю, а среднеквадратическое отклонение — единице.
(ξ — стандартизированная, если $M(\xi) = 0;\ \sigma = 1$.)

любая случайная величина может быть приведена к стандартизованной случайной величине η по формуле: $\eta = \dfrac{\xi - M(\xi)}{\sigma}$. это преобразование включает центрирование случайной величины ($\xi - M(\xi)$) и нормирование (отношение $\dfrac{\xi}{\sigma}$). распределение стандартизованной нормальной случайной величины η называется стандартным нормальным распределением $N(0,\ 1)$ с плотностью
$p(x) = \dfrac{1}{\sqrt{2\pi}} \exp\left(\dfrac{-x^2}{2}\right)$.
(чтобы понять, что это за тупая функция плотности, надо почитать про нормальное распределение, это 4 вопрос из 3 части.)

доп инфа из билетов Левы

```
я хз че это такое.
```

пусть ξ — стандартизированная случайная величина.
тогда:

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-x^2}{2}}$$

пусть $\eta = \sigma\xi + a$.

1. $M(\xi) = \int\limits_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-x^2}{2}}\,dx = 0$
   $M(\eta) = M(\sigma\xi + a) = a$
2. $D(\xi) = M(\xi^2) - (\underbrace{M(\xi)}_{=0})^2 = \int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} x^2 \cdot e^{\frac{-x^2}{2}}\,dx =$

   $\left| \begin{matrix} u = x & du = dx \\ xe^{\frac{-x^2}{2}}\,dx = dv & v = -e^{\frac{-x^2}{2}} \end{matrix} \right|$

   $= \frac{1}{\sqrt{2\pi}} \cdot \left((-xe^{\frac{-x^2}{2}})\Big|_{-\infty}^{\infty} + \int\limits_{-\infty}^{+\infty} e^{\frac{-x^2}{2}}\,dx \right) = 1$ и мне уже пофиг почему.
   $D(\eta) = D(\sigma\xi + a) = \sigma^2 D(\xi) = \sigma^2 = 1$

1 равномерное распределение

```
в первой части уже был вопрос про равномерное
распределение, но немножко другой: про равномерное
распределение на множестве чисел, а это уже подходит
только для дискретной случайной величины.
```

для дискретной случайной величины

ℵ случайная величина ξ имеет равномерное распределение на множестве $\{1, \ldots, n\}$, если вероятность всех значений равна $\frac{1}{n}$.

$$P(\xi = k) = \frac{1}{n}, \quad \text{где } k = 1, \ldots, n$$

$$M(\xi) = \frac{n+1}{2}$$

для непрерывной случайной величины

ξ равномерно распределена на $[a, b]$, если ее плотность $p(x)$ равна

$$p(x) = \begin{cases} \dfrac{1}{b-a}, & \text{если } a \le x \le b, \\ 0, & \text{если } x < a \text{ или } x > b \end{cases}$$

в таком случае функция распределения имеет вид:

$$F(x) = \begin{cases} 0, & \text{если } x < a, \\ \dfrac{x-a}{b-a}, & \text{если } a \le x \le b, \\ 1, & \text{если } x > b. \end{cases}$$

как это получается:

$$F_\xi(x) = \int_{-\infty}^{a} 0\,dx + \int_{a}^{x} \frac{1}{b-a}\,dx = \frac{x-a}{b-a}$$

$$M(\xi) = \frac{a+b}{2}$$

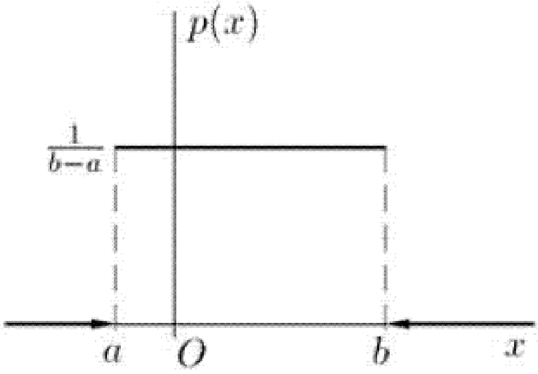


Рис. 6

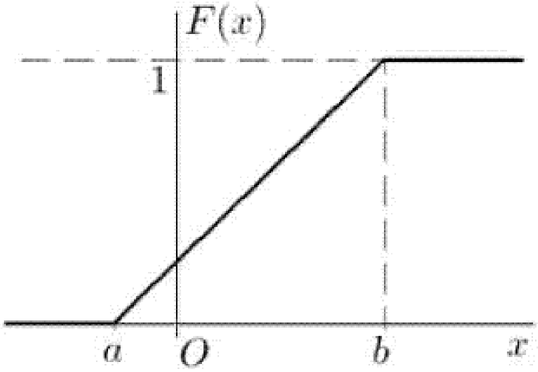


Рис. 7

2 показательное распределение (или экспоненциальное)

ℵ случайная величина распределена по показательному (или экспоненциальному) закону с параметром $\lambda > 0$:

$$p(x) = \begin{cases} 0, & \text{если } x < 0 \\ \lambda e^{-\lambda x} & \text{если } x \geq 0 \end{cases}$$

$$F(x) = \begin{cases} 0, & \text{если } x < 0 \\ 1 - e^{-\lambda x} & \text{если } x \geq 0 \end{cases}$$

$$M(\xi) = \frac{1}{\lambda}$$

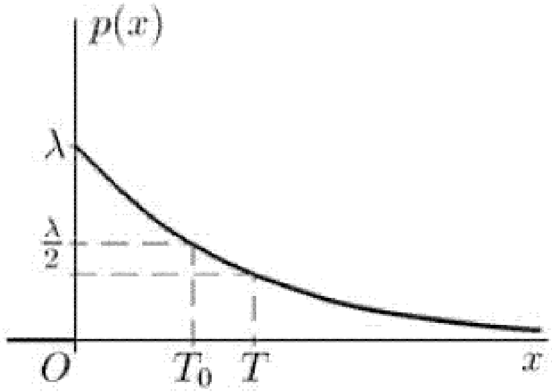


Рис. 8

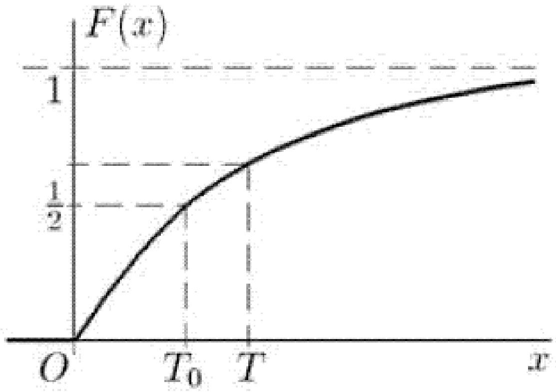


Рис. 9

экспоненциально распределенная случайная величина может принимать только положительные значения.

3 характеристическое свойство показательного распределения

характеристическое свойство показательного (экспоненциального) распределения — отсутствие памяти.

пусть ξ распределена по экспоненциальному закону.

$A = \{\xi > t\}$
$B = \{\xi > t + \tau\}$

$$P({}^{B}|_{A}) = P({}^{\xi > t+\tau}|_{\xi > t}) = \frac{P(\xi > t + \tau)}{P(\xi > t)} = \frac{e^{-\lambda(t+\tau)}}{e^{-\lambda t}} = e^{-\lambda\tau} = P(\xi > \tau)$$

расшифруем, что это значит

предположим, у нас есть какое-то устройство, и мы знаем, какова вероятность того, что оно прослужит t времени и $t + \tau$ времени. оказывается, что вероятность того, что устройство прослужит $t + \tau$ времени, когда оно уже прослужило t времени, равна вероятности того, что новое устройство прослужит τ времени. такая аналогия, кстати, работает для устройств с полупроводниками, потому что их время работы распределено экспоненциально.

$\star$ можно сказать, что если случайная величина обладает свойством отсутствия памяти, то она распределена экспоненциально, и наоборот.

4 нормальное распределение

ℵ непрерывная случайная величина η распределена по нормальному (или гауссову) закону $N(a,\ \sigma)$ с параметрами $(a,\ \sigma)$, где $a \in (-\infty;\ +\infty)$, если плотность распределения равна

$$p_\eta(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{\frac{-(x-a)^2}{2\sigma^2}}$$

здесь a — среднее значение, обычно матожидание, а σ — среднее квадратическое отклонение.
чтобы не ломать глаза, $e^{\text{что-нибудь}}$ частенько записывают как $\exp(\text{что-нибудь})$:

$$p_\eta(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\left(\frac{-(x-a)^2}{2\sigma^2}\right)$$

функция распределения для нормального распределения равна

$$F_\eta(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \int\limits_{-\infty}^{x} \exp\left(\frac{-(\tau-a)^2}{2\sigma^2}\right) d\tau$$

```
вся та же фигня, только с интегралом и т вместо x, потому
что производная экспоненты и все дела.
```

держите картинку:

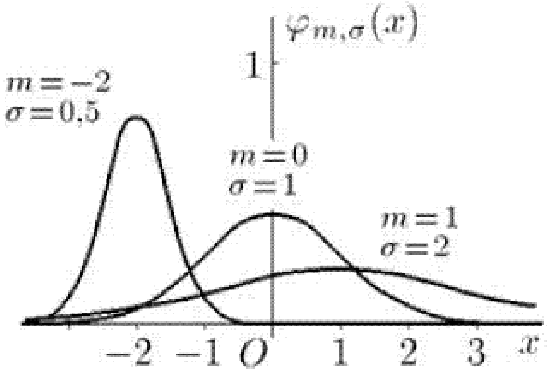


Рис. 10

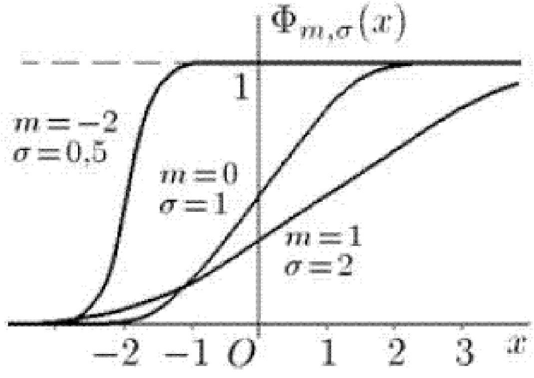


Рис. 11

(здесь $m = a$, $\phi(x) = p(x)$, $\Phi(x) = F(x)$.)

как видно из рисунков, параметр a (на графике m) определяет положение влево-вправо на графике, а параметр σ — насколько "острым" будет пик плотности.

нормальное распределение становится стандартизированным, если $a = 0,\ \sigma = 1.$

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(\frac{-x^2}{2}\right)$$

если выразить нормальное распределение η через стандартизированное ξ (или наоборот), то получится, что

$$\xi = \frac{\eta - a}{\sigma}; \quad \eta = \sigma\xi + a$$

$F_\eta(x)$ — интеграл, не берущийся в квадратуру (нельзя выразить через элементарные функции / взять интеграл), поэтому мы выражаем его через F_ξ, которая уже четкий пацанчик и считается по-человечески.

$$F_\eta(x) = P(\eta < x) = P(\sigma\xi + a < x) = P(\xi < \frac{x-a}{\sigma}) = F_\xi(\frac{x-a}{\sigma})$$

функция Лапласа:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x \exp\left(\frac{-\tau^2}{2}\right) d\tau$$

$$F_\xi(x) = \frac{1}{2} + \Phi(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int\limits_0^x \exp\left(\frac{-\tau^2}{2}\right) d\tau$$

```
объяснение, почему интеграл от 0 до x, от госпожи Марии:
Функция это интеграл от -бесконечности до икса. Поскольку
распределение стандартизованное, то график плотности
симметричен относительно 0. То есть интеграл от
-бесконечности до 0 равен интегралу от 0 до +бесконечности,
а поскольку площадь под всем графиком это 1, то каждый из
этих кусков равен 1/2. Теперь к собсна функции: хз зачем,
возможно для удобства, всю функцию поделили на кусок от
-бесконечности до 0 и кусок от 0 до икса. Первый кусок это
1/2, как сказано выше, а второй это Ф(x).
```

матожидание стандартизованного распределения равно 0, а обычное получается подстановкой замены:

$M_\xi = \int\limits_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = (\text{внесение } x \text{ под } dx) =$
$-\frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2}\right) d\left(-\frac{x^2}{2}\right) = 0$

тогда:

$$M_\eta = M(\ \sigma\xi + a\ ) = M(\ \sigma\xi\ ) + M(a) = 0 + a = a$$

с дисперсией та же история, но дисперсия стандартизованного равна 1:

$D(\xi) = M(\xi^2) - (\underbrace{M(\xi)}_{=0})^2 = \int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} x^2 \cdot e^{\frac{-x^2}{2}}\ dx =$

$\begin{vmatrix} u = x & du = dx \\ xe^{\frac{-x^2}{2}}\ dx = dv & v = -e^{\frac{-x^2}{2}} \end{vmatrix}$

$= \frac{1}{\sqrt{2\pi}} \cdot \left((-xe^{\frac{-x^2}{2}})\Big|_{-\infty}^{\infty} + \int\limits_{-\infty}^{+\infty} e^{\frac{-x^2}{2}}\ dx \right) = 1$ и мне уже кристалльно пофиг, почему.

$$D(\eta) = D(\sigma\xi + a) = \sigma^2 D(\xi) = \sigma^2 = 1$$

5 интеграл Лапласа и его связь со случайной величиной, распределенной по нормальному закону

нормальное распределение становится стандартизированным, если $a = 0, \ \sigma = 1.$

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(\frac{-x^2}{2}\right)$$

если выразить нормальное распределение η через стандартизированное ξ (или наоборот), то получится, что

$$\xi = \frac{\eta - a}{\sigma}; \quad \eta = \sigma\xi + a$$

$F_\eta(x)$ — интеграл, не берущийся в квадратуру (нельзя выразить через элементарные функции / взять интеграл), поэтому мы выражаем его через F_ξ , которая уже четкий пацанчик и считается по-человечески.

$$F_\eta(x) = P(\eta < x) = P(\sigma\xi + a < x) = P(\xi < \frac{x-a}{\sigma}) = F_\xi(\frac{x-a}{\sigma})$$

функция Лапласа:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x \exp\left(\frac{-\tau^2}{2}\right) d\tau$$

$$F_\xi(x) = \frac{1}{2} + \Phi(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int\limits_0^x \exp\left(\frac{-\tau^2}{2}\right) d\tau$$

```
объяснение, почему интеграл от 0 до x, от госпожи Марии:
Функция это интеграл от -бесконечности до икса. Поскольку
распределение стандартизованное, то график плотности
симметричен относительно 0. То есть интеграл от
-бесконечности до 0 равен интегралу от 0 до +бесконечности,
а поскольку площадь под всем графиком это 1, то каждый из
этих кусков равен 1/2. Теперь к собсна функции: хз зачем,
возможно для удобства, всю функцию поделили на кусок от
-бесконечности до 0 и кусок от 0 до икса. Первый кусок это
1/2, как сказано выше, а второй это Ф(x).
```

6 функция от случайной величины и формулы для вычисления ее характеристик

пусть $(\Omega,\ \mathfrak{A},\ P)$ — вероятностное пространство, ξ — случайная величина и $y = f(x)$ — некоторая функция.

рассмотрим $\eta = f(\xi)$:
если для любого $x : \{\omega : \eta < x\} = \{\omega : f(\xi) < x\}$, то η — тоже случайная величина.

функция от дискретной случайной величины

пусть дискретная ξ задана законом распределения: $\{x_i,\ P_i\}_{i \in I}$. тогда $\eta = f(\xi)$ — тоже дискретная случайная величина.

при этом $\eta = \{f(x_i),\ P_i'\}$ задается значениями функции и соответствующими им вероятностями.
если некоторое значение функции можно получить из нескольких значений аргументов, то вероятность этого значения функции складывается из вероятностей всех подходящих для нее аргументов.

ξ	-1	0	1
P	0.5	0.4	0.1

$\eta = \xi^2$	0	1
P	0.4	0.6

функция от непрерывной случайной величины

пусть непрерывная ξ задана плотностью $p_\xi(x)$ и функцией распределения $F_\xi(x)$; $\eta = f(\xi)$.

если $f(x)$ монотонно возрастает, то она обратимая:

$$F_\eta(x) = P(\eta < x) = P(f(\xi) < x) = P(\xi < f^{-1}(x)) = F_\xi(f^{-1}(x)) \quad (1)$$

$$p_\eta(x) = \frac{dF_\eta(x)}{dx} = \frac{d(F_\xi(f^{-1}(x)))}{dx} = \frac{dF_\xi(f^{-1}(x))}{dx} \cdot \frac{df^{-1}(x)}{dx} \quad (2)$$

если $f(x)$ монотонно убывает, то:

$$F_\eta(x) = P(f(\xi) < x) = P(\xi \geq f^{-1}(x)) = 1 - P(\xi < f^{-1}(x)) =$$

$$= 1 - F_\xi(f^{-1}(x)) \tag{3}$$

$$p_\eta(x) = \frac{dF_\eta(x)}{dx} = \frac{d(1 - F_\xi(f^{-1}(x)))}{dx} = -\frac{d(F_\xi(f^{-1}(x)))}{dx} \cdot \frac{df^{-1}(x)}{dx}$$

обобщим (2) и (4):

$$p_\eta(x) = \frac{d(F_\xi(f^{-1}(x)))}{dx} \cdot \left|\frac{df^{-1}(x)}{dx}\right| = p_\xi(f^{-1}(x)) \cdot \left|\frac{df^{-1}(x)}{dx}\right| \tag{5}$$

```
вот этого якобы не было на лекциях, переписала у Левы. мне
не совсем понятно, чего тут происходит, поэтому не могу
гарантировать происходящее.
```

если $f(x)$ не монотонна, то:

1. разбиваем на участки монотонности $\Delta i(y)$, где $f(x) < y$
2. $$F_\eta(y) = \sum_{i \in I} \int\limits_{\Delta i(y)} p_\xi(x)\,dx; \quad p_\eta(y) = F'_\eta(y)$$

числовые характеристики

$$M(\eta) = \begin{cases} \sum\limits_{i \in I} f(x_i) P_i, & \text{если } \xi \text{ дискретная,} \\ \int\limits_{-\infty}^{\infty} f(x) p_\xi(x) dx, & \text{если } \xi \text{ непрерывная.} \end{cases}$$

$$D(\eta) = \begin{cases} \sum\limits_{i \in I} (f(x_i) - M_\eta)^2 P_i, & \text{если } \xi \text{ дискретная,} \\ \int\limits_{-\infty}^{\infty} (f(x_i) - M_\eta)^2 p_\xi(x) dx, & \text{если } \xi \text{ непрерывная.} \end{cases}$$

7 логарифмически нормальное распределение

говорят, что η имеет логарифмически нормальное распределение, если
$\xi = \ln \eta$ распределен по нормальному закону.
$\eta > 0; \quad \eta = e^{\xi}$
$\ln \eta \sim N(a,\ \sigma)$

ξ распределена нормально:

$$p_{\xi}(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\left(\frac{-(x-a)^2}{2\sigma^2}\right)$$

тогда:

$$p_{\eta}(x) = \begin{cases} 0, & x \leq 0, \\ \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\left(\frac{-(\ln x - a)^2}{2\sigma^2}\right) \cdot \frac{1}{x}, & x > 0. \end{cases}$$

8 формула для вычисления математического ожидания и дисперсии для случайной величины Y, являющейся функцией от случайной величины X

$$M(\eta) = \begin{cases} \sum\limits_{i \in I} f(x_i) P_i, & \text{если } \xi \text{ дискретная,} \\ \int\limits_{-\infty}^{\infty} f(x) p_\xi(x) dx, & \text{если } \xi \text{ непрерывная.} \end{cases}$$

$$D(\eta) = \begin{cases} \sum\limits_{i \in I} (f(x_i) - M_\eta)^2 P_i, & \text{если } \xi \text{ дискретная,} \\ \int\limits_{-\infty}^{\infty} (f(x_i) - M_\eta)^2 p_\xi(x) dx, & \text{если } \xi \text{ непрерывная.} \end{cases}$$

9 многомерная дискретная случайная величина

```
то же, что билет 1 из 2 части. кстати, там дальше будет
много повторов, поэтому я просто буду называть их часть-
номер. типа, билет 2-1 - 2 часть 1 билет.
```

пусть на одном и том же вероятностном пространстве $(\Omega, \mathfrak{A}, P)$ задано n случайных величин: $\xi_1 = \xi_1(\omega), \dots, \xi_n = \xi_n(\omega)$.

ℵ упорядоченный набор случайных величин $(\xi_1, \dots, \xi_n)$ называется многомерной (n-мерной) случайной величиной или случайным вектором. ξ_i $(i \in I)$ — компонента случайного вектора.

функция распределения:
для одномерной случайной величины $F_\xi(x) = P(\xi < x)$
для многомерной: $F_{\xi_1 \dots \xi_n}(x_1, \dots, x_n) = P\underbrace{(\xi_1 < x_1, \dots, \xi_n < x_n)}_{\in \mathfrak{A}}$ — функция распределения n-мерного случайного вектора (также называют совместной функцией распределения случайных величин $\xi_1, \dots, \xi_n$).

ℵ n-мерная случайная величина распределена по дискретному закону, если $(X_i, P_i)_{i \in I}$, где $X_i = (x_i^1, \dots, x_i^n)$, $P_i = P(\xi = X_i)$ — закон распределения дискретной случайной величины (т. е. все ξ_i — дискретные случайные величины.)

для двумерной случайной величины:
$i = 1, \dots, n$, $j = 1, \dots, m$

$$\sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} = 1$$

10 непрерывная многомерная случайная величина

пусть на одном и том же вероятностном пространстве $(\Omega,\ \mathfrak{A},\ P)$ задано n случайных величин: $\xi_1 = \xi_1(\omega),\ \ldots,\ \xi_n = \xi_n(\omega)$.

ℵ упорядоченный набор случайных величин $(\xi_1,\ \ldots,\ \xi_n)$ называется многомерной (n-мерной) случайной величиной или случайным вектором. $\xi_i\ (i \in I)$ — компонента случайного вектора.

функция распределения:
для одномерной случайной величины $F_\xi(x) = P(\xi < x)$
для многомерной: $F_{\xi_1 \ldots \xi_n}(x_1,\ \ldots,\ x_n) = P\underbrace{(\xi_1 < x_1,\ \ldots,\ \xi_n < x_n)}_{\in \mathfrak{A}}$ — функция распределения n-мерного случайного вектора (также называют совместной функцией распределения случайных величин $\xi_1,\ \ldots,\ \xi_n$).

ℵ непрерывной двумерной случайной величиной $(\xi,\ \eta)$ называется такая двумерная случайная величина, функция распределения $F_{\xi,\,\eta}(x_1,\ x_2) = P(\xi < x_1,\ \eta < x_2)$ которой может быть представлена в виде:

$$F_{\xi,\,\eta}(x_1,\ x_2) = \iint\limits_{\substack{-\infty < y_1 < x_1 \\ -\infty < y_2 < x_2}} p(y_1,\ y_2)\, dy_1\, dy_2$$

это — двукратный интеграл по области $\{y_1 < x_1,\ \ldots,\ y_n < x_n\}$. по теореме Фубини этот интеграл можно представить в виде повторного, причем в любом порядке:

$$F_{\xi,\,\eta}(x_1,\ x_2) = \int\limits_{-\infty}^{x_1} dy_1 \int\limits_{-\infty}^{x_2} p(y_1,\ y_2) dy_2 = \int\limits_{-\infty}^{x_2} dy_2 \int\limits_{-\infty}^{x_1} p(y_1,\ y_2) dy_1$$

$p(y_1,\ y_2)$ называется совместной плотностью распределения случайных величин ξ и η.

будем предполагать, что $p(x_1,\ x_2)$ — непрерывная.
тогда:

$$p(x_1,\ x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1,\ x_2) = \frac{\partial^2}{\partial x_2 \partial x_1} F(x_1,\ x_2)$$

свойства совместной плотности

1. $p(x_1,\ x_2) \geq 0$
2. $P(a_1 < \xi < b_1,\ a_2 < \eta < b_2) = \int\limits_{a_1}^{b_1} dx_1 \int\limits_{a_2}^{b_2} p(x_1,\ x_2)\, dx_2$
3. $\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} p(x_1,\ x_2)\, dx_1\, dx_2 = 1$
4. $P(x_1 < \xi < x_2 + \Delta_1,\ x_2 < \eta < x_2 + \Delta_2) \approx p(x_1,\ x_2)\Delta_1\Delta_2$
5. $P(\xi = x_1,\ \eta = x_2) = 0$
6. $P((\xi,\ n) \in D) = \iint\limits_D p(x_1,\ x_2)\, dx_1\, dx_2$

Пусть D — некоторая область на плоскости (рис. 4).

Тогда, как следует из свойства 4, вероятность попадания двумерной случайной величины (ξ, η) в малый прямоугольник $\{a < \xi_1 < a + \Delta_1,\ b < x_2 < b + \Delta_2\}$ приближенно равна $p(a, b)\Delta_1\Delta_2$. Поскольку попадания в непересекающиеся прямоугольники являются несовместными событиями, то, для того чтобы найти полную вероятность попадания двумерной случайной величины (ξ, η) в область D, нужно просуммировать вероятности попадания во все «малые» прямоугольники, входящие в область D. Переходя к пределу, получаем для $\mathsf{P}\{(\xi, \eta) \in D\}$ — вероятности попадания (ξ, η) в область D — формулу

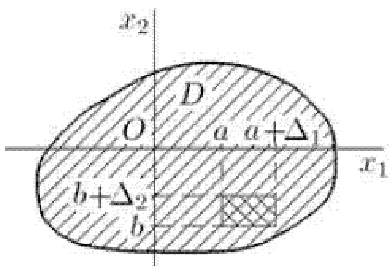


Рис. 4

$$\textbf{6.}\ \mathsf{P}\{(\xi, \eta) \in D\} = \iint\limits_D p(x_1, x_2)\, dx_1\, dx_2.$$

7. $p_\xi(x) = \int\limits_{-\infty}^{\infty} p_{\xi,\eta}(x,\ y)\, dy, \quad p_\eta(x) = \int\limits_{-\infty}^{\infty} p_{\xi,\eta}(x,\ y)\, dx$

доказательство:

Далее, из свойства 7 совместной функции распределения и определения совместной плотности распределения имеем

$$F_\xi(x) = F_{\xi,\eta}(x,\infty) = \int\limits_{-\infty}^{x} dy_1 \int\limits_{-\infty}^{\infty} p_{\xi,\eta}(y_1,y_2)\ dy_2,$$

$$F_\eta(x) = F_{\xi,\eta}(\infty,x) = \int\limits_{-\infty}^{x} dy_2 \int\limits_{-\infty}^{\infty} p_{\xi,\eta}(y_1,y_2)\ dy_1,$$

откуда, дифференцируя по x, получаем выражения для одномерных плотностей распределения случайных величин ξ и η

11 свойства двумерной функции распределения

ℵ функцией распределения двумерного случайного вектора $\xi = (\xi_1,\ \xi_2)$ называется функция $F_\xi(x_1,\ x_2) = P(\xi_1 < x_1,\ \xi_2 < x_2)$. альтернативное название: совместная функция распределения случайных величин $\xi_1,\ \xi_2$.

свойства:

1. $0 \leq F(x_1,\ x_2) \leq 1$
2. $F(x_1,\ x_2)$ — неубывающая по каждому из двух аргументов.
3. $F(-\infty,\ x_2) = F(x_1,\ -\infty) = 0$, т. к. $\{\xi_1 < -\infty\}$ и $\{\xi_2 < -\infty\}$ — невозможные события.
4. $F(\infty,\ \infty) = 1$
5. $P(a_1 \leq \xi_1 < b_1,\ a_2 \leq \xi_2 < b_2) = F(a_1,\ a_2) + F(b_1,\ b_2) - F(b_1, a_2) - F(a_1,\ b_2)$

4. $F(\infty, \infty) = 1$.

Найдем вероятность попадания двумерной случайной величины (ξ_1, ξ_2) в прямоугольник $\{a_1 \leqslant x_1 < b_1,\ a_2 \leqslant x_2 < b_2\}$ (рис. 2). Для этого сначала определим вероятность попадания в полуполосу $\{x_1 < b_1,\ a_2 \leqslant x_2 < b_2\}$. Но эта вероятность представляет собой вероятность попадания в квадрант $\{x_1 < b_1,\ x_2 < b_2\}$ за вычетом вероятности попадания в квадрант $\{x_1 < b_1,\ x_2 < a_2\}$, т. е.

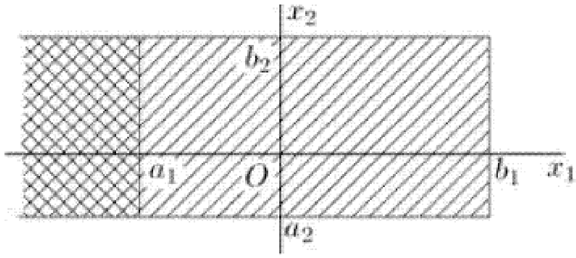


Рис. 2

$$\mathsf{P}\{\xi_1 < b_1,\ a_2 \leqslant \xi_2 < b_2\} = F(b_1, b_2) - F(b_1, a_2).$$

Теперь осталось заметить, что вероятность попадания в прямоугольник $\{a_1 \leqslant x_1 < b_1,\ a_2 \leqslant x_2 < b_2\}$ есть вероятность попадания в полуполосу $\{x_1 < b_1,\ a_2 \leqslant x_2 < b_2\}$, из которой вычтена вероятность попадания в полуполосу $\{x_1 < a_1,\ a_2 \leqslant x_2 < b_2\}$, равная $F(a_1, b_2) - F(a_1, a_2)$. Окончательно получаем

5. $\mathsf{P}\{a_1 \leqslant \xi_1 < b_1,\ a_2 \leqslant \xi_2 < b_2\} = F(b_1, b_2) - F(b_1, a_2) - F(a_1, b_2) + F(a_1, a_2)$.

6. $F(x_1,\ x_2)$ — непрерывная слева по каждому из аргументов.
7. $F_{\xi_1, \xi_2}(x,\ \infty) = F_{\xi_1}, \quad F_{\xi_1, \xi_2}(\ \infty,\ x) = F_{\xi_2}$ — условие согласованности.

Наконец, последнее свойство устанавливает естественную связь между распределением случайного вектора (ξ_1, ξ_2) и распределениями случайных величин ξ_1 и ξ_2. Событие $\{\xi_2 < \infty\}$ достоверно, поэтому $\{\xi_1 < x_1\} \cap \{\xi_2 < \infty\} = \{\xi_1 < x_1\}$. Аналогично $\{\xi_1 < \infty\} \cap \{\xi_2 < x_2\} = \{\xi_2 < x_2\}$. Значит,

7. $F_{\xi_1,\xi_2}(x,\infty) = F_{\xi_1}(x), \;\; F_{\xi_1,\xi_2}(\infty,x) = F_{\xi_2}(x).$

прикольный факт: это маргинальные функции распределения.

$$F_\xi(x_1, +\infty) = F_{\xi_1}(x_1)$$
$$F_\xi(+\infty, x_2) = F_{\xi_2}(x_2)$$

12 свойства совместной плотности распределения

1. $p(x_1,\ x_2) \geq 0$
2. $P(a_1 < \xi < b_1,\ a_2 < \eta < b_2) = \int\limits_{a_1}^{b_1} dx_1 \int\limits_{a_2}^{b_2} p(x_1,\ x_2)\, dx_2$
3. $\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} p(x_1,\ x_2)\, dx_1\, dx_2 = 1$
4. $P(x_1 < \xi < x_2 + \Delta_1,\ x_2 < \eta < x_2 + \Delta_2) \approx p(x_1,\ x_2)\Delta_1\Delta_2$
5. $P(\xi = x_1,\ \eta = x_2) = 0$
6. $P((\xi,\ n) \in D) = \iint\limits_D p(x_1,\ x_2)\, dx_1\, dx_2$

Пусть D — некоторая область на плоскости (рис. 4).

Тогда, как следует из свойства 4, вероятность попадания двумерной случайной величины (ξ, η) в малый прямоугольник $\{a < \xi_1 < a + \Delta_1,$ $b < x_2 < b + \Delta_2\}$ приближенно равна $p(a, b)\Delta_1\Delta_2$. Поскольку попадания в непересекающиеся прямоугольники являются несовместными событиями, то, для того чтобы найти полную вероятность попадания двумерной случайной величины (ξ, η) в область D, нужно просуммировать вероятности попадания во все «малые» прямоугольники, входящие в область D. Переходя к пределу, получаем для $\mathsf{P}\{(\xi, \eta) \in D\}$ — вероятности попадания (ξ, η) в область D — формулу

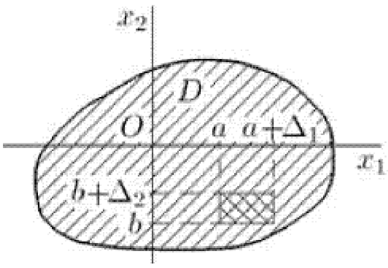


Рис. 4

6. $\mathsf{P}\{(\xi, \eta) \in D\} = \iint\limits_D p(x_1, x_2)\, dx_1\, dx_2.$

7. $p_\xi(x) = \int\limits_{-\infty}^{\infty} p_{\xi,\,\eta}(x,\ y)\, dy, \quad p_\eta(x) = \int\limits_{-\infty}^{\infty} p_{\xi,\,\eta}(x,\ y)\, dx$

доказательство:

Далее, из свойства 7 совместной функции распределения и определения совместной плотности распределения имеем

$$F_\xi(x) = F_{\xi,\eta}(x, \infty) = \int\limits_{-\infty}^{x} dy_1 \int\limits_{-\infty}^{\infty} p_{\xi,\eta}(y_1, y_2)\, dy_2,$$

$$F_\eta(x) = F_{\xi,\eta}(\infty, x) = \int\limits_{-\infty}^{x} dy_2 \int\limits_{-\infty}^{\infty} p_{\xi,\eta}(y_1, y_2)\, dy_1,$$

откуда, дифференцируя по x, получаем выражения для одномерных плотностей распределения случайных величин ξ и η

13 маргинальная плотность, функция распределения отдельных компонент

в теории вероятности и статистике, маргинальное распределение подмножества набора случайных величин — это распределение вероятностей переменных, содержащихся в этом подмножестве. это даёт возможность представить вероятности различных значений переменных в подмножестве без указания на другие значения переменных.

термин маргинальная величина используется для обозначения переменных из подмножества, в котором они содержатся. эти термины называют «маргинальными», потому что они были найдены путем суммирования значений в таблице вдоль строк или столбцов, и записи этой суммы в полях таблицы.

Случай двух переменных

	x_1	x_2	x_3	x_4	$p_y(Y)\downarrow$
y_1	$4/32$	$2/32$	$1/32$	$1/32$	$8/32$
y_2	$2/32$	$4/32$	$1/32$	$1/32$	$8/32$
y_3	$2/32$	$2/32$	$2/32$	$2/32$	$8/32$
y_4	$8/32$	0	0	0	$8/32$
$p_x(X) \rightarrow$	$16/32$	$8/32$	$4/32$	$4/32$	$32/32$

Совместное и маргинальное распределение случайных величин X,Y имеют ненулевую взаимную информацию I(X; Y). Значения совместного распределения представляют собой квадрат 4×4, значения маргинального распределения указаны на полях справа и снизу.

маргинальная плотность распределения

(7 св-во из 12 билета)

$$p_\xi(x) = \int\limits_{-\infty}^{\infty} p_{\xi,\,\eta}(x,\ y)\, dy$$

$$p_\eta(x) = \int\limits_{-\infty}^{\infty} p_{\xi,\,\eta}(x,\ y)\, dx$$

маргинальная функция распределения

(7 св-во из 11 билета)

$$F_\xi(x_1, +\infty) = F_{\xi_1}(x_1)$$

$$F_\xi(+\infty, x_2) = F_{\xi_2}(x_2)$$

14 начальные и центральные моменты многомерного распределения

начальный момент $\alpha_{k,s}$

дискретная случайная величина

$$\alpha_{k,s} = \sum_{i \in I} \sum_{j \in J} x_i^k y_i^s \cdot P_{ij}$$

непрерывная случайная величина

$$\alpha_{k,s} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^k y^s \cdot p_{\xi\eta}(x, y)\, dxdy$$

центральный момент $\mu_{k,s}$

дискретная случайная величина

$$\mu_{k,s} = \sum_{i \in I} \sum_{j \in J} (x_i - M(\xi))^k \cdot (y_j - M(\eta))^s \cdot P_{ij}$$

непрерывная случайная величина

$$\mu_{k,s} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - M(\xi))^k \cdot (y - M(\eta))^s \cdot p_{\xi\eta}(x, y)\, dxdy$$

частные случаи:

- $\alpha_{1,0} = M_\xi$
- $\alpha_{0,1} = M_\eta$
- $\mu_{2,0} = D_\xi$
- $\mu_{0,2} = D_\eta$
- $\mu_{1,1} = cov(\xi, \eta)$ — смешанный момент порядка 1, 1

15 ковариационная и корреляционная матрицы случайного вектора

```
и какого черта этого не было на лекциях?
```

ковариационная матрица случайного вектора

[предлагаю почитать википедию.](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%B2%D0%B0%D1%80%D0%B8%D0%B0%D1%86%D0%B8%D0%BE%D0%BD%D0%BD%D0%B0%D1%8F_%D0%BC%D0%B0%D1%82%D1%80%D0%B8%D1%86%D0%B0)
(https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%B2%D0%B0%D1%80%D0%B8%D0%B0%D1%86%D0%B8%D0%BE%D0%BD%D0%BD%D0%B0%D1%8F_%D0%BC%D0%B0%D1%82%D1%80%D0%B8%D1%86%D0%B0)

ℵ ковариационная матрица (или матрица ковариаций) — это матрица, составленная из попарных ковариаций элементов одного или двух случайных векторов.

ковариационная матрица случайного вектора — квадратная симметрическая неотрицательно определенная матрица, на диагонали которой располагаются дисперсии компонент вектора, а внедиагональные элементы — ковариации между компонентами.

дана многомерная случайная величина $(\xi_1, \ldots, \xi_n)$.

$$cov(\xi_i,\ \xi_j) = M[(\xi_i - M(\xi_i))(\xi_j - M(\xi_j))]$$

составим матрицу из ковариаций:

$$k = \begin{pmatrix} cov(\xi_1,\ \xi_1) & cov(\xi_1,\ \xi_2) & \ldots & cov(\xi_1,\ \xi_n) \\ cov(\xi_2,\ \xi_1) & cov(\xi_2,\ \xi_2) & \ldots & cov(\xi_2,\ \xi_n) \\ \ldots & \ldots & \ldots & \ldots \\ cov(\xi_n,\ \xi_1) & cov(\xi_n,\ \xi_2) & \ldots & cov(\xi_n,\ \xi_n) \end{pmatrix}$$

зачем это надо? никто не знает.

5a. $$\operatorname{cov}(\xi,\xi) = M(\xi\xi^T) - M\xi M\xi^T =$$

$$= \begin{pmatrix} M\xi_1^2 - (M\xi_1)^2 & \dots & M(\xi_1\xi_n) - M\xi_1 M\xi_n \\ \dots & \dots & \dots \\ M(\xi_n\xi_1) - M\xi_n M\xi_1 & \dots & M\xi_n^2 - (M\xi_n)^2 \end{pmatrix} =$$

$$= \begin{pmatrix} D\xi_1 & \dots & \operatorname{cov}(\xi_1,\xi_n) \\ \dots & \dots & \dots \\ \operatorname{cov}(\xi_n,\xi_1) & \dots & D\xi_n \end{pmatrix} = \begin{pmatrix} \sigma_{11}^2 & \dots & \sigma_{1n}^2 \\ \dots & \dots & \dots \\ \sigma_{n1}^2 & \dots & \sigma_{nn}^2 \end{pmatrix}.$$

Дисперсии координат можно представить в виде

$$\sigma_{11} = \sigma_1^2, \quad \sigma_{22} = \sigma_2^2, \dots, \quad \sigma_{kk} = \sigma_k^2,$$

корреляционная матрица случайного вектора

$$\rho = \frac{cov(\xi_i,\ \xi_j)}{\sigma_i\sigma_j}$$

корреляционная матрица:

$$r = \begin{pmatrix} 1 & \rho(\xi_1,\ \xi_2) & \dots & \rho(\xi_1,\ \xi_n) \\ \rho(\xi_2,\ \xi_1) & 1 & \dots & \rho(\xi_2,\ \xi_n) \\ \dots & \dots & \dots & \dots \\ \rho(\xi_n,\ \xi_1) & \rho(\xi_n,\ \xi_2) & \dots & 1 \end{pmatrix}$$

единицы на главной диагонали, потому что $\rho(\xi_i,\ \xi_i) = \frac{cov(\xi_i,\ \xi_i)}{\sigma_i\sigma_i} =$ ну там надо пораскладывать, получить из матожидания дисперсию, из дисперсии σ, и тогда получится $= \frac{\sigma_i\sigma_i}{\sigma_i\sigma_i} = 1$.

и вообще, корреляционная матрица — результат умножения вектора $(\xi - M(\xi))$ на себя же, но транспонированного. типа $(\xi - M(\xi)) \cdot (\xi - M(\xi))^T$.

16 если модуль коэффициента корреляции равен 1, то случайные величины линейно зависимы

```
огромнейшее спасибо Марии за это доказательство.
```

пусть $\zeta = a\xi + b\eta$ (да-да, дзета. потому што буквы кончелись.)

- $M_\zeta = M(a\xi) + M(b\eta) = aM_\xi + bM_\eta$
  (по какому-то из свойств матожидания, и нам не нужно, чтобы ξ и η были независимыми.)
- $M_\zeta^2 = a^2 M_\xi^2 + 2abM_\xi M_\eta + b^2 M_\eta^2$
  (квадрат суммы.)
- $M(\zeta^2) = M(a^2\xi^2 + 2ab\xi\eta + b^2\eta^2) = a^2M(\xi^2) + 2abM(\xi\eta) + b^2M(\eta^2)$
  (опять квадрат суммы, но в скобочках и раскрываем матожидание.)
- $D_\zeta = M(\zeta^2) - M_\zeta^2 = a^2M(\xi^2) - a^2M_\xi^2 + 2abM(\xi\eta) - 2abM_\xi M_\eta + b^2M(\eta^2) - b^2M_\eta^2$
  $= a^2(M(\xi^2) - M_\xi^2) + 2ab(M(\xi\eta) - M_\xi M_\eta) + b^2(M(\eta^2) - M_\eta^2) =$
  $a^2D_\xi + 2ab\, cov(\xi,\ \eta) + b^2 D_\eta =$

  $$\rho(\xi,\ \eta) = \pm 1 \Rightarrow \frac{cov(\xi,\ \eta)}{\sigma_\xi \sigma_\eta} = \pm 1 \Rightarrow cov(\xi,\ \eta) = \pm\sigma_\xi\sigma_\eta$$

- $= a^2\sigma_\xi^2 \pm 2ab\, \sigma_\xi\sigma_\eta + b^2\sigma_\eta^2 = (a\sigma_\xi \pm b\sigma_\eta)^2$

то есть $D_\zeta = (a\sigma_\xi \pm b\sigma_\eta)^2$.

т. к. $a,\ b \in \mathbb{R} \Rightarrow$ всегда существуют такие $a,\ b$, что $D_\zeta = 0$.

$D_\zeta = 0 \Rightarrow \zeta = const$.

пусть $\zeta = c$.

$a\xi + b\eta = c \Rightarrow \xi = \dfrac{c - b\eta}{a} \Rightarrow \xi$ и η линейно зависимые, что и требовалось доказать.

в обратную сторону тоже работает

```
и доказывается за три секунды. и Рогов сказал, что надо доказывать блин.
```

пусть ξ и η — линейно зависимые, то есть $\eta = a\xi + b$. найдем коэффициент корреляции:

$$\rho(\xi,\ \eta) = \frac{cov(\xi,\ \eta)}{\sigma_\xi \sigma_\eta}$$

$$cov(\xi,\ \eta) = M(\xi\eta) - M(\xi)M(\eta)$$
$$\sigma = \sqrt{D}$$

получаеца:

$$\rho(\xi,\ \eta) = \frac{M(\xi(a\xi + b)) - M(\xi)M(a\xi + b)}{\sqrt{D(\xi)} \cdot \sqrt{D(a\xi + b)}} = \frac{aD(\xi)}{\sqrt{D(\xi)}\sqrt{a^2 D(\xi)}} =$$

$$= \frac{aD(\xi)}{|a|D(\xi)} = \begin{cases} 1, & a > 0 \\ -1, & a > 0 \end{cases}$$

то есть $\rho(\xi,\ \eta) = \pm 1$, что и требовалось доказать.

17 неравенство Маркова

пусть ξ - положительно определённая случайная величина с конечным матожиданием. тогда:

$$\forall \varepsilon > 0 : \quad P(\xi \geq \varepsilon) \leq \frac{M(\xi)}{\varepsilon}$$

доказательство:

рассмотрим событие $A = \{\xi \geq \varepsilon\}$. введём индикатор события (билет 2-5) A: 0, если событие не наступило и 1, если наступило:

$$I_A = \begin{cases} 0, & \overline{A} : 1 - P(A) \\ 1, & A : P(A) \end{cases} \quad \Rightarrow M(I_A) = P(A)$$

$$\varepsilon \cdot I_A \leq \xi \cdot I_A \leq \xi$$

опустим среднюю часть неравенства и найдем мат. ожидание крайних:

$$M(\varepsilon I_A) \leq M_\xi$$

$M(\varepsilon) = \varepsilon$ (константа), $M(I_A) = P(A)$.

$$\varepsilon \cdot P(A) \leq M(\xi) \Big/ : \varepsilon$$

$$P(A) \leq \frac{M(\xi)}{\varepsilon}$$

аналогично доказывается следствие:

$$P(\ |\xi| \geq \varepsilon\ ) \leq \frac{M(\xi^k)}{\varepsilon^k}$$

```
(как — аналогично — я пока не придумала, кушойте что есть.)
```

Пример

Пусть студенты никогда не приходят вовремя, они всегда опаздывают. В среднем они опаздывают на 3 минуты. Какова вероятность того, что студент опоздает на 15 минут и более? Дать грубую оценку сверху.

$$\mathbb{P}(|\xi| \geqslant 15) \leqslant \frac{3}{15} = 0.2$$

18 неравенство Чебышева

```
простите, я уже просто переписываю — слабо понимаю, что происходит.
```

для каждой случайной величины ξ, имеющей дисперсию $D(\xi)$, при любом $\varepsilon > 0$ справедливо неравенство:

$$P(|\xi - M(\xi)| \geq \varepsilon) \leq \frac{D(\xi)}{\varepsilon^2}$$

доказательство:

следствие неравенства Маркова: $P(\ |\xi| \geq \varepsilon\ ) \leq \frac{M(\xi^k)}{\varepsilon^k}$

пусть $\xi = |\xi - M(\xi)|; \quad k = 2$. подставим в вот ту ↑ формулу.

$$P(|\xi - M(\xi)| \geq \varepsilon) \leq \frac{M[(\xi - M(\xi))^2]}{\varepsilon^2} = \frac{D(\xi)}{\varepsilon^2}$$

следствие 1: $P(|\xi - M(\xi)| < \varepsilon) \geq 1 - \frac{D(\xi)}{\varepsilon^2}$

доказательство:

"величина больше или равна ε" и "величина меньше ε" — противоположные события.

следствие 2: пусть $\xi_1, \xi_2, \dots, \xi_n$ — независимые случайные величины с конечной дисперсией

$$P\left(\left|\frac{\xi_1 + \dots + \xi_n}{n} - \frac{M(\xi_1) + \dots + M(\xi_n)}{n}\right| < \varepsilon\right) \geq 1 - \frac{1}{n\varepsilon^2} \cdot \frac{D_{\xi_1} + \dots + D_{\xi_n}}{n}$$

доказательство: следствие следствия 1 если $\xi = \frac{\xi_1 + \dots + \xi_n}{n}$

тогда общее матожидание — это среднее арифметическое всех матожиданий.

аналогично $D(\xi) = \frac{1}{n^2}(D(1) + \ \dots \ + D(n))$

19 закон больших чисел для одинаково распределенных случайных величин

это ряд теорем о сходимости среднего арифметического случайных величин к среднему арифметическому их ожиданий.

закон больших чисел можно записать в виде неравенства Чебышева:
пусть $\xi_1, \ \ldots, \xi_n$ — последовательность независимых случайных величин, для которых существует такое c, что $\forall \, i: \ D(\xi_i) \leq c$ (т. е. дисперсия ограничена сверху числом c). тогда для любого $\varepsilon > 0$ выполняется тождество:

$$\lim_{n\to\infty} P\left(\left|\frac{\xi_1 + \ \ldots \ + \xi_n}{n} - \frac{M(\xi_1) + \ \ldots \ + M(\xi_n)}{n}\right| < \varepsilon\right) = 1$$

эта формула означает, что при увеличении n среднее арифметическое случайных величин сходится к среднему арифметическому их матожиданий.

ну такая теорема, логичная что ли.

доказательство

$$P\left(\left|\frac{\xi_1 + \ \ldots \ + \xi_n}{n} - \frac{M(\xi_1) + \ \ldots \ + M(\xi_n)}{n}\right| < \varepsilon\right) \leq 1$$

(по св-ву вероятности...)

по следствию 2 из неравенства Чебышева:

$$P(\text{че там выше было}) \geq 1 - \frac{1}{n\varepsilon^2} \cdot \frac{D(\xi_1) + \ \ldots \ + D(\xi_n)}{n} \geq 1 - \frac{1}{n\varepsilon^2} \cdot c$$

т. к. $\lim\limits_{n\to\infty} 1 - \dfrac{c}{n\varepsilon^2} = 1$, то $1 \leq \lim\limits_{n\to\infty} P(\ldots) \leq 1$. по теореме о зажатой переменной $\lim\limits_{n\to\infty} P(\ldots) = 1$.

следствие - закон больших чисел для одинаково распределенных случайных величин

пусть $\xi_1, \ \ldots, \ \xi_n$ — последовательность независимых одинаково распределенных случайных величин. т. к. они одинаково распределены, то все их числовые характеристики, в частности, матожидание и дисперсия, равны между собой.

эта последовательность сходится к числу a по вероятности, если для любого $\varepsilon > 0$ верно тождество:

$$\lim_{n\to\infty} P\left(\left|\frac{\xi_1 + \ \dots \ + \xi_n}{n} - a\right| < \varepsilon\right) = 1$$

если n достаточно большое, то матожидание каждого распределения можно приравнять к среднему арифметическому всех случайных величин в последовательности.

20 теорема Бернулли

теорема Бернулли — частный случай закона больших чисел.

пусть имеется схема Бернулли: n — количество испытаний, p — вероятность успеха. пусть индикаторы $I_1, \ \ldots, \ I_n$ содержат в себе 0 и 1 — результаты испытаний. $\mu(n) = \sum\limits_{i=1}^{n} I_i$ — количество успешных испытаний.

тогда для любого $\varepsilon > 0$ выполняется тождество:

$$\lim_{n \to \infty} P\left(\left|\frac{\mu(n)}{n} - p\right| < \varepsilon\right) = 1$$

```
здесь просто написано, что с увеличением количества
испытаний процент успешных опытов стремится к вероятности
успеха. эта теорема с математической точки зрения
объясняет, почему точность вычисления вероятности какого-то
события увеличивается, если мы проводим больше опытов.
```